Universiteit
Leiden
The Netherlands

Mongolic phonology and the Qinghai-Gansu languages

Nugteren, H.

Citation
Nugteren, H. (2011, December 7). *Mongolic phonology and the Qinghai-Gansu languages*. *LOT dissertation series*. Utrecht : LOT, Netherlands Graduate School of Linguistics. Retrieved from https://hdl.handle.net/1887/18188

Version:	Not Applicable (or Unknown)
License:	Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden
Downloaded from:	https://hdl.handle.net/1887/18188

Note: To cite this publication please use the final published version (if applicable).

3. MODERN DEVELOPMENT OF THE CM VOWELS

3.1. Introduction

In the following pages the main developments of each CM vowel will be discussed. Each section will start with the 'default' development, which need not be the most frequent development. The quality of unaccented vowels is rather unstable in the QG languages, and easily influenced by the consonant environment. After the default reflexes the most common conditioned changes will be discussed. Whenever possible the focus will be on correspondences of historical and comparative importance.

In all modern languages the CM vowels underwent several changes, which are correlated to changes in the vowel system of each language as a whole. Such changes often affect the number of vowel phonemes, and modify or undermine vowel harmony. On the level of the lexeme such changes affect both the quantity and quality of the vowels.

In all three peripheries we find that the original vocalism is simplified. In general the distinction between **ï* and **i* is (all but) absent, and the number of phonemic rounded vowel qualities was reduced. The Dagur system and the Shirongol system are the result of different routes of simplification. In Dagur the front rounded vowels merged into *u*, the back ones into *ɔ*.[100] In Shirongol (Kangjia excepted) **ö* and **ü* merged with their respective back counterpart.

Loss of productive vowel harmony in the Shirongol languages removed one of the factors that tend to block or slow down vowel shifts, as well as restrict the changes in individual words.

Before discussing the modern developments of each CM vowel phoneme, some general remarks are due. Although most vowel shifts in the QG languages do not constitute autonomous sound laws, some patterns and tendencies can be recognized. Apart from the phonetic properties of the vowels themselves, such as vowel height, there are several factors influencing the vowel development in a given word in a given language, including (former or extant) vowel harmonic constraints, accent, and phonetic environment, i.e. the place or manner of articulation of the adjacent consonants, and syllable structure.

Vowel harmony

As mentioned above, the vowel system was simplified in the QG languages. However, even in Baoan, where all active vowel harmony has been lost, its role in shaping words is still evident in much of the surviving Mongolic lexicon.

The vowel harmony of the ancestral stages of Eastern Yugur and Shirongol may have been unlike the CM system. Eastern Yugur suggests it goes back to a Turkic-type front-back system. Monguoric may have had a simplified system of vowel qualities *a e i o u* without productive harmony. Kangjia suggests a system in

[100] This does not apply in all environments. If the following vowel is long *a:* both **o* and **u* result in Dag *u*, e.g. *tuwa:* < **togaan* 'cauldron', *xula:n* < **hulaan* 'red'.

which the front vowels tended to be centralised, while the other Baoanic languages are uninformative as the front rounded vowels have been lost.

In CM there must have been suffixes with the alternations *A* (= **a* or **e*), *U* (= **u* or **ü*), and *I* (=**ï* or **i*). The non-high vowel suffixes (with *A*) later developed additional variants with **o* and **ö*, but high-vowel suffixes with fourfold alternation never arose.

The Shirongol languages preserve traces of palatal and labial harmony in existing formations only. Word shapes in Mongghul and Dongxiang show that they once had suffixes with the alternation *a/e/o* (a fourth alternant cannot be demonstrated). Within Shirongol only Kangjia roughly preserves the distinction in word stems between 'front' *ö* and *ü* and 'back' *o* and *u*; the distinction is no longer relevant in suffixation.

Eastern Yugur retains most of palatal harmony and labial harmony in productive suffixes (except in the original short high vowels). Vowel changes, mostly under the influence of the consonants, may lead to historically unexpected (unharmonic) vowel sequences within stems, as in EYu *neyan* < **nayan* 'eighty'. Nevertheless the accented vowel will continue to trigger the (also diachronically) regular suffix shapes.

Harmonic rounding of the non-first *a/e* in words like **olan* 'many' and **köken* 'chest' was already in progress in Middle Mongol. It is therefore not surprising that both Eastern Yugur and Shirongol feature this rounding. Only Dongxiang has several exceptions, e.g. *nokie* 'friend', as opposed to EYu *nökör*, MgrH *nokor*, BaoD *nokor*, Kgj *nəχgʁ* < **nöker*. As Dongxiang is clearly embedded in the Baoanic branch of Shirongol, it must for now be assumed that forms like *nokie* are due to secondary unrounding.[101]

Accent

The development of the vowels is most predictable in monosyllabic stems, and in the accented vowel of longer stems. In Dagur, the accent falls on the first vowel. In Shirongol it falls on the last vowel, whereas accent placement in Eastern Yugur is more complicated, and partly correlated with vowel length (see 3.14.5.), as in Khalkha and elsewhere in Central Mongolic. A less rigid accent placement may also have existed in premodern Shirongol, judging from modern word shapes.

Phonetic stability of unaccented vowels and their phonological distinctiveness are low. In Dagur this typically applies to all non-first vowels, in Shirongol to all non-last vowels. Apart from changing the shape of individual lexemes, and the phonemic status of their component parts, the instability of unaccented vowels also tends to obscure the historical information we are interested in here.

The unaccented vowels are unstable in that they are likely to be reduced, and then inclined to adopt articulation features from the phonetic environment, mostly from the surrounding consonants, but also from the accented vowel. The

[101] Secondary unrounding also took place in Oirat-Kalmuck, as hinted at by Rybatzki 2003b:368. That this is indeed the case is seen most clearly in words like *jɔla:* < **jɔlɔ:* < **jïlɔ:* < **jïloa* 'rein'.

following examples from Mongghul show some of the possible influences of the consonants on the unaccented vowel:

Mongghul	CM	
ala-	**ala-*	to kill
xara	**kara*	black
ara:l ~ ra:l	**aral*	river
bara:- ~ bura:-	**bara-*	to finish
ʒ́æġa ~ ʒ́iġa	**ǰaka*	collar

Unaccented vowels may even be elided, provided that the resulting word structure is allowed. Through such elisions, the differences in accent have created some of the more dramatic differences between Mongolic languages, as in the following examples:

Baoan	Dagur	CM	
la-	*wail-*	**uïla-*	to cry
sə	*ɔs*	**usun*	water
duŋ ~ ʂduŋ	*šid*	**sidün*	tooth

Such cases also demonstrate the complementary role the languages have in reconstruction.

Vowel height

High vowels are most susceptible to changes caused by adjacent consonants, reduction, and elision. However, unaccented non-high vowels are often raised in the process of reduction, and are then likely to undergo similar developments.

In Eastern Yugur and Monguoric all four simple high vowels **ï *i *u *ü* often merge into *ə*, even in the accented syllable.[102] Although *ə* is phonetically no longer a high vowel, it will not be confused with the original non-high vowels, as these retain a full pronunciation when carrying the accent.

Elision of an unaccented vowel, whether originally high or non-high, is possible when the resulting sequence of consonants is permitted in the language in question.

Non-high simple vowels may be lengthened. Such lengthening appears to be mostly related to accent and syllable structure. Not all vowel lengths (that do not stem from contractions of double vowel sequences) can at this stage be explained.

In short, the developments that a non-high vowel may undergo are more diverse, and which development they will undergo in any individual word is harder to predict.

[102] Most words that do not undergo this development contain another high rounded vowel, like EYu *nʉdʉn*, MgrH *nudu* < **nidün* ‘eye’. This can be viewed as a kind of assimilation unrelated to vowel harmony.

3.1.1. Proto Mongolic legacy and prehistoric shifts

Some changes in the vocalism need not, and in some cases, should not be attributed to the modern languages. Several developments found in today's QG languages had already started in Middle Mongol.

Harmonic rounding is well represented in Middle Mongol, albeit not consistently applied.

There are also some instances in Middle Mongol of the famous regressive assimilation phenomenon known as palatal breaking (see 3.6.). Cases of assimilation of diphthongs, e.g. **ïa* > **aa*, are also documented early.

The distinction between **ï* and **i* may have been lost and redeveloped several times in different periods and regions.[103] It is worth noting that even within the QG languages both the separate status of the vowels **ï* and **i* and their merger into a single high unrounded vowel **i* is reflected. In the stage leading up to the Monguoric subgroup of Shirongol **ï* and **i* apparently merged into a palatal vowel that palatalised preceding **k-* > *ć-* both in words of both original vowel classes. In the ancestral stage of the Baoanic languages the two vowels remained different enough to cause the phoneme **k-* to split into a velar **k-* and a uvular **q-*.[104] A similar split occurred in Eastern Yugur.[105]

None of these early changes were completed in the Middle Mongolic period. Palatal breaking has been a tendency for centuries without becoming universally applied (e.g. Dagur preserves unbroken **ï*/**i* in many native words, but features breaking even in some loanwords from Tungusic). Harmonic rounding of non-first syllables is reversible, as can be seen occasionally in Dongxiang (and as a rule in Kalmuck).

3.1.2. Types of changes

The default development of accented vowels is usually quite straightforward. The tendency of **ï*/**i* to be 'neutralized' into *ə* is found in all QG languages. The 'front' rounded vowels **ö* and **ü* have merged with their 'back' counterparts in all of Shirongol except Kangjia. Apart from Kangjia, the Shirongol languages typically give the impression that they go back to an ancestral stage with a five-vowel system *a e i o u*. However, the fact that Kangjia preserves four rounded vowel phonemes makes it impossible to reconstruct the five-vowel system for Baoanic, even less so for Proto Shirongol.

Most other kinds of vowel development in the QG languages involve assimilation and/or reduction. The latter may also lead to elision. Non-high vowels

---

[103] This still holds true if one departs from an original CM system of seven vowels.

[104] The vowels**ï* and **i* themselves do not survive in Eastern Yugur or any Baoanic language. It is true that Dongxiang typically has the pronunciation *ï* following *q-* and *ġ-*, but *ï* may have (re)developed in recent centuries due to the uvular consonants.

[105] In Eastern Yugur the status of the phonemes */k/* and */q/* is, at least in the analysis of Bolčuluu, less strong, because only **ï* and **i* merged into *ə*, while *ö* and *ü* did not merge with their back counterparts. Eastern Yugur thus lacks oppositions like Mangghuer *qo-* 'to dry' versus *ko-* 'to swell' (cf. **koa-* and **köe-*).

can undergo lengthening. The main types that will be repeatedly encountered are the following.

Assimilation phenomena

Assimilation of first syllable**ï* and **i* by the second syllable vowel (‘palatal breaking’):
CM **čïsun* > Dgx *ɕusuŋ* ‘blood’
Assimilation of the middle vowel of trisyllables:
CM **čimegen* > MgrH *ćimuge* ‘marrow’ (due to the *-m-*)
CM **kïtuga* > **kïtoga* > Dgx *qudoğo* ‘knife’ (due to the *a* of the following syllable)

Palatalisation by the preceding consonant:
CM **čagaan* > Kgj *čiχɔ ~ čɔχɔ* ‘white’
Palatalisation by **ï*/**i* of the next syllable (umlaut):
CM **talbï-* > MgrM *tebi-* ‘to put’

Harmonic rounding by the vowel of the preceding syllable:
CM **kola* > Dgx *ġolo* ‘far’
Rounding of first syllable vowel by the vowel of the second
CM **kalaun* > BaoX *χoloŋ* ‘hot’ (via **χaloŋ*)
Labialisation by adjacent consonant
CM **marta-* > MgrH *muṣda:-* ‘to forget’
Delabialisation by adjacent consonant:
CM **bučalga-* > EYu *pəǰalğa-* ‘to boil (tr.)’

Reduction and elision phenomena

Raising of (unaccented) non-high vowels:
CM **ora-* > MgrH *uro-* ‘to enter’
Raising and subsequent palatalisation:
CM **noyan* > EYu *niyɔn* ‘lord’
Elision of unaccented vowels:
CM **kočar-* > BaoÑ *hʒ́ər-* ‘to remain’

Lengthening phenomena

Lengthening of accented last vowel:
CM **untara-* > MgrH *ntəra:-* ‘to sleep’
Lengthening before high vowel of the final syllable:
CM **morïn* > EYu *mɔ:rə* ‘horse’

Non-first vowels in the North

Due to the different accent placement the Mongolic subgroups complement one another in reconstruction.

In Dagur and Central Mongolic the accented vowel of the first syllable is most stable, except **ï* and **i* which have almost disappeared due to the phenomenon of '*i*-breaking', which involves the 'copying' of the second vowel into the first vowel position and the reduction of the **ï* and **i* originally occupying that position. The other vowels of the first syllable may be palatalised (to various degrees) if an **ï* or **i* follows later in the word.

Short vowels of the non-first syllables are often altered and/or reduced. In Dagur and the three Central Mongolic standard languages the short vowels of non-first syllables are quite limited.[106] Khalkha only allows either *i* or 'non-*i*', which takes the shape of a reduced non-high vowel the quality of which can be predicted from the environment.

Dagur allows *ə*, *i*, and *u* as non-first vowels (especially word-finally, the latter two tend to be reduced to ^{y} and ^{w}, i.e. palatalisation and labialisation of the preceding consonant).[107] This means that in Dagur there are three distinctive short vowels in non-first positions, each of which can occur in words of either harmonic class. Although in many cases Dagur *ə* in these positions stems from CM **a* or **e*, *i* from **ï* and **i*, and *u* from **u* and **ü*, the Dagur vowels do not necessarily reflect historical reality as expected. Dag *alim* 'pear' and *am*^{y} 'life' correctly reflect the **ï* of CM **alïma*, **amïn*, and Dag *taryun* 'fat' and *alk*^{w}- 'to step' correctly reflect the **u* of CM **targun*, **alku-*. But in other words the same *i/y* and *u/w* are simply products of the environment. The appearance of *i* instead of *ə* may be triggered by a palatal consonant, as in *gaǰir* 'land' < **gaǰar, əǰin* 'master' < **eǰen*. Dagur *u* appears (instead of the normally expected *ə*) after a rounded first vowel or a diphthong ending with a rounded second element, as in *auɣun* 'firstborn' < **aukan, nuɣ*^{w} 'hole' < **nüken*. As a consequence, Dagur forms whose vocalism is in agreement with a known ancient form, such as *uɣ*^{w}- < **ükü-* 'to die', do not constitute independent evidence for this CM reconstruction.

Non-last vowels in the South

In the QG languages it is the vowel of the last syllable that is typically most stable, while vowels of the remaining syllables may be influenced by the phonetic environment. Unaccented vowels often develop a different pronunciation under the influence of the consonantal environment or of the vowels of the following syllable(s). They are susceptible to devoicing, reduction, or even elision. The changes in the non-first syllables are too unpredictable to enable us to define a subsystem of reduced vowels, as is possible in the non-first syllables of Dagur.

---

[106] Not all dialects that belong to Central Mongolic share these same restrictions. Khamnigan and Ordos do distinguish between short *A*, *U*, and (neutral) *i* in non-first syllables. Therefore this treatment of the vocalism of the non-first syllables can not be considered a feature of the shared ancestor of the central languages.

[107] Svantesson et al. analyse this differently (2003:150). They see *u* of non-first syllables as the 'shwa phoneme' coloured by a preceding labialised consonant, and *i* of non-first syllables as the shwa phoneme coloured by a preceding palatalised consonant. Here I will use an analysis and notation that seeks to find a middle ground between indicating the pronunciation correctly and make the history of each word as transparent as possible.

Reduction in QG typically involves neutralisation of non-high vowels into *ə*[108], after which they are often raised and palatalised > *i*, or raised and labialised > *u*, mainly depending on the articulation of the preceding consonant. Secondary developments such as MgrH *ćiġa:n* 'white', *ćiʑaġ* 'flower', *ʑil* 'year', *ʑila:* 'lamp', *ʑira:-* 'to mix', *ʑirge* 'heart' (from **čagaan*, **čečeg*, **ǰil*, **ǰula*, **ǰuura-*, **ǰürgen*) illustrate how the distinctive value of the first syllable has diminished in some phonetic environments.[109]

Elision of vowels

There are two general routes for elision, which cannot necessarily be distinguished in individual lexemes. The first is simply the loss of an unaccented, often already reduced, vowel. The second is through devoicing of the vowel, typically when preceding a strong consonant. Elision through either route takes place frequently if the resulting consonant configuration is permitted. The general 'guideline' applying to both Dagur and the QG languages seems to be that all unaccented vowels can in principle be reduced, and elided as long as this does not result in the creation of inadmissible consonant sequences. E.g. CM **anda* 'friend' may become Dag *and*, and **arasun* 'skin' may become Dag *ars*, but **aral* 'shaft' may not become Dag **arl* (the actual Dagur form is now *arəl ~ allə*). CM **halagan* 'palm of the hand' became MgrH *xalġa*, but **xlaġa* (or **xlġa*) would not be admissible sequences, although we do have *nʑasə* 'plough' from **anǰasun*, *ʂde* 'early' from **erte*[110] as these sequences happen to be permitted. The number and character of permitted initial consonant clusters varies from dialect to dialect.[111] In QG, elisions are most frequent in Eastern Yugur, Mongghul and Baoan, as these are most tolerant of consonant clusters. But even languages with similar restrictions may opt for different solutions, as in the case of CM **sïnaga*, which has produced trisyllabic modern forms such as EYu *šənağa*, MgrH *śinaġa*, and Dgx *ʂïnağa*, as well as forms with reduced first or second vowel, such as MgrH *ʂnaġa*, MgrM *ʂəŋġa*, BaoÑ *naġa*.

---

[108] In QG there is also a strong tendency to merge the high vowels into *ə*. The fact that reduction of the non-high vowels does not normally take place in accented syllables, ensures that sufficient distinctiveness is preserved in the vowel system.

[109] These developments naturally complicate the discussion of palatal breaking in the QG languages (see 3.6.).

[110] Mongghul *xalġa* < **halagan* 'palm of the hand' is related to a tendency for *Mittelsilbenschwund* when it results in permitted consonant clusters. However, low vowels are less likely to disappear. Clusters *rg/rġ* and *lg/lġ* seem to especially 'popular', so the middle vowel is lost in all following words, creating disyllabic stems in Mongghul: **heligen* 'liver', **kulagaï* 'thief', **malagaï* 'hat', **solagaï* 'left', **tolagaï* 'head', **berigen* 'daughter-in-law', **gölige* 'puppy', **silükei* 'saliva', **herike* 'beads', **kurïgan* 'lamb', **kürigen* 'brother-in-law', **korakaï* 'insect'. Conversely, some other clusters tend to be broken up by epenthetic vowels, e.g. *ʂdoġo-* 'to prick' < **kadku-*, *ʂdoġo:n* 'dense' < **ödken*.

[111] Obviously the rules for permissible consonant clusters may change over time within each language. This is most notable in the (Sanchuan and) Mangghuer data through the decades. Mangghuer developed a form *muba-* 'to swim' from the form *mba-* with an initial cluster it apparently inherited but no longer allowed (ultimately < **humba-*).

Since elision is possible but not obligatory, the modern forms are not entirely predictable from the CM forms. Not only is the above general rule used differently by each language, we also find variation within each language, as in MgrH *arasə ~ rasə*, EYu *arasən ~ arəsən ~ arsən*, BaoD *arsuŋ*, Dgx *arasuŋ* < **arasun* 'skin'.[112]

Another route whereby vowels can be lost in Eastern Yugur and Shirongol is devoicing. This occurs when a non-last vowel (vowel height is irrelevant in this case) is followed by a strong consonant, or located between strong consonants.

Even though the elisions can be found in several QG languages, the treatment of individual stems is very idiosyncratic. They must therefore be seen as having taken place independently, and can not be reconstructed for an ancestral language. We will in the following simplify matters, and assume that vowel loss preceding a weak consonant took place by reduction, and that vowel loss preceding a strong consonant is related to devoicing.

Dongxiang and Moghol, which also have word-final accent, are most conservative in that they usually preserve the vowels of all original syllables of longer words (related to their low tolerance of consonant clusters).

Although in QG the non-first vowels are more stable and probably more important phonologically, as well as more useful for historical comparison, they can be influenced by the vowel of the first syllable. In MgrH *yo:-* 'to sew' (< CM **oya-*) the *o* is due to the rounded first syllable which was subsequently lost. Likewise in EYu *hkor* 'bovine' (< CM ***hüker***) the rounded vowel is due to the first vowel that was lost at a later stage.

The accented syllables play an important phonological role in keeping similar-sounding lexemes apart. Even in these non-standardised languages, free variation is less abundant in the accented vowel. While variations such as EYu *čağa:n ~ čəğa:n ~ čğa:n* 'white' (**čağaan*), *ǰağasən ~ ǰəğasən* 'fish' (**ǰagasun*) are very common and meaningless, the same vowels are less likely to vary in the accented syllable, as in EYu *šərə-* 'to stitch' ≠ *šəra-* 'to roast', *hsə-* 'to go sour' ≠ *hsa-* 'to make', *čüsə-* 'to cut' ≠ *čüsa-* 'to make'.[113]

Vowel lengthening

Under various circumstances short vowels have been lengthened, both in Dagur and in the QG languages. Such developments will be discussed in 3.14.

3.1.3. Vowel detection

As in the central languages, even the original number and location of the vowels within the stem is not always obvious, in that the modern form may not reflect the original situation. Often the place of the vowels is synchronically determined by

[112] This diversity means that the reductions and elisions can generally not be used in reconstruction forms.

[113] The first example continues an old opposition **sïrï-* 'to stitch' ≠ **sïra-* 'to roast', but phonemic oppositions of this kind need not be inherited from CM, as the other examples demonstrate. For the other Eastern Yugur verbs see **is-* 'to go sour', **ǰasa-* 'to make', **ǰisü-* 'to cut', and again **ǰasa-* 'to make' in the comparative supplement.

permissible consonant sequences rather than by etymology. Compare the following words, which behave similarly in modern Dagur, but originally had different structures: *mələɣ* ‘frog’ (< CM **melekei*), genitive *məlɣiː*, versus *tərəɣ* ‘cart’ (< **tergen*), genitive *tərɣiː*. The first word was allowed to lose its medial vowel in the genitive, whereas the second added a vowel in the nominative to break up its original cluster. Modern rules block clusters like *lɣ* and *rɣ* in final position but the same sequence is allowed medially.

Another peculiarity Dagur shares with central Mongolic is ‘palatal creep’, whereby an **i* (from **ï*/*i*) in a non-first syllable is carried over to an earlier or later position in the word, as in Dag *adilʲ* < **adalï* ‘similar’, *ariɣʲ* < **arakï* ‘liquor’, *alli-* < **arïl-* ‘to clean’. Such cases illustrate how both the vowel quality and the exact position of the vowels in the word would be difficult to reconstruct based on one Mongolic subgroup alone.

Vowel detection problems in the QG languages also usually involve the breaking up of consonant clusters, e.g. EYu *qutul-* < **hogtal-* ‘to fell’, MgrH *dəge-* < **hedke-* ‘to cut’.

3.2. Development of simple **a*

3.2.1. Default development

In monosyllables CM **a* has usually been preserved as such. Its quality is rarely influenced by surrounding consonants. For monosyllables ending in a vowel see 3.14.3.

In monosyllables ending in a consonant both quality and quantity are mostly left intact in Dagur and the QG languages. The same normally applies to **a* in accented syllables of longer stems. In Dagur, as in central Mongolic, the accent normally falls on the first syllable. In the QG languages **a* in the accented final syllable is preserved, while the vowels of non-last syllables are susceptible to change, reduction and elision. However, in words with certain syllable structures not conducive to elision, and in the vicinity of 'neutral' consonants unlikely to influence vowel quality, **a* is preserved even in the unaccented syllables. Neutral consonants are those who do not palatalise or labialise the adjacent vowels, i.e. *d, t, g, k, s, l, n, ŋ, r*. Elision will often not take place in words without a strong consonant in medial position. Words with unreduced *a*'s in most languages include **ala* 'crotch', **alda-* 'to lose', **aman* 'mouth', **arga* 'way', **harban* 'ten', **kabar* 'nose', *kada* 'rock', **kara* 'black', **naran* 'sun', **tala* 'plain'. Examples:

CM	**gal* 'fire'[114]	**gar* 'hand'	**sam* 'comb'[115]	**ala-* 'kill'	**sara* 'moon'
Dag	*gal^{y}*	*gar^{y}*	*sannə, sand*	*al-*	*sar*
EYu	*ġal*	*ġar*	*sam*	*ala-*	*sara*
MgrH	*ġal*	*ġar*	*sam, san*	*ala-*	*sara*
MgrM	*ġar*	*ġar*	*saŋ*	*ala-*	*sara*
BaoD	*χal*	*χar*	*sam*	*alə-*	*sarə*
BaoÑ	*χal*	*χar*	*sam*	*ala-*	*sara*
Kgj	*χar*	*χar*	*san, saɾ̥*	*ala-*	*sara, cara*
Dgx	*qaŋ*	*qa*	*saŋ*	*ala-*	*sara*
Mog	*ɣo:l*	*ɣar*	---	*ola-*	---

Accented **a* tends to be preserved, not only after **a* of the first syllable, but also after **ï* (e.g. **sïra* 'yellow'); after **u* (**kura* 'rain'); after **aa, *aï,* and **au* (**daara-* 'to feel cold', **naïman* 'eight', **ǰaura* 'space between').

Exceptions may occasionally arise when monosyllabic verbs gain a final vowel in the Shirongol languages, and thus develop like primarily disyllabic stems. The added vowel usually seems to stem from the connective vowel *u* needed to connect the stem to some suffixes, which was then reinterpreted as part of the stem. In Dongxiang, this metanalysis may in part be motivated by a desire to maintain the integrity of the stem, since the final consonant would not be permitted in that

[114] The palatal element in Dagur *gal^{y}* and *gar^{y}* at first sight suggests a disyllabic origin, but the stem shapes may have been adopted from inflected forms such as the genitive case *gali:, gari:*.

[115] The Dagur forms and Kangjia *saɾ̥* seem to be due to a false analysis of the derived verbs **samda-* and **samla-* 'to comb', repectively.

position; the alternative would be to drop the offending final consonant altogether.[116] The verb **gar-* 'to come out' thus gained a final vowel in Dongxiang and some Baoan dialects, thus shifting the accent away from the **a*, resulting in Dgx *qïri-*, BaoD *χărə-*. Compare Dgx *qa* 'hand' from the, originally homophonous, noun **gar*, where a similar solution to preserve the final consonant was not present.[117] In the case of CM **čad-* 'to be satiated', the unexpected form *ɕudu-* in both Mangghuer and Dongxiang is apparently caused by the connective vowel *u* which was incorporated into the stem.

3.2.2. Harmonic rounding

Most Mongolic languages feature some degree of rounding harmony, causing **a* of the non-first syllable to become *o* after **o* of the first syllable. This is an old development already documented in Middle Mongol, and common in the central languages, as well as in QG. In Eastern Yugur we tend to get **a* > *ɔ*, as in *χɔlɔ* < **kola* 'far', *dɔğɔlɔŋ* < **dogalaŋ* 'limping'. In Shirongol the sequence *o - a* has developed into *o - o* (with the same diphthongisations as the *o* that stems from primary **o/*ö*), e.g. BaoD *soġor* (~ *soχur* with raising) < **sokar* 'blind', *ġor* < **hokar* 'short', *ġolo* < **kola* 'far'.[118] In Kangjia both *ɔ*'s tend to be raised in words of this structure, e.g. *ġulu* < **kola* 'far', *uru-* < **ora-* 'to enter'.[119] Other stems with similar developments in Shirongol include **olaŋ* 'bellyband', **oŋgača* 'trough', **sogta-* 'to get drunk', **solaŋga* 'rainbow'.

In Eastern Yugur and both Monguor languages the same rounding can usually be found after a complex vowel containing **o*, e.g. EYu *χɔinɔ*, MgrM *quaino /qoino/* < **koïna* 'posterior'. Dongxiang preserves **a* in these words, e.g. *quaina* < **koïna*.[120]

In Dagur we see two groups of reflexes, neither of which seems to reflect harmonic rounding. The first group features labial breaking, which shows that **a* must still have been present (preserved, or possibly redeveloped) in the second syllable at the time this breaking took place.

---

[116] In nouns final **r* is generally dropped, however short the resulting form.

[117] It may be relevant that noun stems are used in their shortest form more often than verb stems, which can only be used as such in the function of basic imperative form.

[118] Although many of these cases of rounding may be very early, some rounding must have happened more recently in stems that originally did not contain **o*, e.g. BaoD *doġo* ~ *doġə* < **kïtuga* 'knife', *solġo* < **saulga* 'bucket', MgrH *ʑo:ro* < **ʃaura* 'space between'.

[119] This is also applied to loanwords, e.g. Kgj *suχu* 'Mongolian'< Amdo *sok-kwa*, lit. Tib *sog-pa*.

[120] The reflexes after **oa* are less predictable. Within Mongghul we find alternations such as *to:la- ~ to:lo-* < **toa-la-* 'to count'. In Eastern Yugur *du:ra* < **doara* 'underside', the raising of **oa* > *o:* > *u:* blocked the harmonic rounding (cf. Mongghul *do:ro*).

Dagur	CM	
dʷatər	**dotar*	inside
xʷakər	**hokar*	short
walən	**olan*	many
war-	**ora-*	to enter

In the second group the second vowel has disappeared altogether, e.g. *ɔi-* < **oya-* 'to sew', *xɔl* < **kola* 'far', *sɔns-* < **sonas-* 'to hear', or it has developed into *-u-*, when syllable structure prevents elision, e.g. *bɔɣunʸ* < **boganï* 'low', *čɔlpun* < **čolban* 'Venus', *sɔɣur* < **sokar* 'blind'. This *-u-* does not count as evidence for the existence of harmonic rounding in earlier Dagur. It is not clear how the division between the two groups of words with the different developments came about.[121]

3.2.3. Reduction/raising

Reduction of unaccented **a* is common in most Mongolic languages, as is its elision in certain environments (see below). In Dagur these phenomena happen in non-first syllables, as in Dag *xaləɣ* < **halagan* 'palm of the hand'.

Reduction of **a* in the first syllable is a common development in all of QG. Reduction and raising is often accompanied by devoicing when a strong consonant is following and/or preceding the vowel. The primary result of reduction is the vowel *ə*, as in EYu *pəsa* < **basa* 'also', EYu *qəša:* < **kasïa* 'fence', BaoD *χəra* < **kara* 'black', *iəsə-* 'to make', perhaps from local Turkic *yasa-* rather than < **ǰasa-*.

In Dongxiang this reduced vowel *ə* takes the shape *ï* after *q-* and after the retroflex affricates *ʈ-* and *ʈʂ-*, e.g. Dgx *qïʈa* < **kačar* 'cheek', *qïʂaŋ* < **kašaŋ* 'lazy'. This reduction mainly seems to occur between strong consonants, and may typically involve devoicing. In cases such as Dgx *ʈïɡaŋ* 'white', ultimately from CM **čagaan*, it cannot be established whether Dgx *ï* may have developed < *i* < *ə* < *a*, that is, via the secondary *i* found in Baoan and Mongghul.

Reduction of unaccented *a* is the stage preceding labialisation and palatalisation as well as complete loss all of which are discussed below.

As can be seen in *sarə* < **sara* 'moon' above, Dahejia, and to a lesser extent other Baoan dialects, have a tendency to reduce **a* > *ə* even in accented final syllables.[122] This includes cases with harmonic rounding. In view of the equivalents in closely related languages this development is relatively recent. Examples:

---

[121] There are also some etyma which show both of these developments in Dagur, e.g. *dʷatər* 'inside' as opposed to *dətur* 'internal organs'. Enkhbat suggests that the former is from CM **dotara*, the second from **dotar*. In fact in this case one would expect both CM forms to result in the same Dagur shape. Some unexpected Dagur forms may be the result of borrowing from central Mongolic dialects, Manchu, or Northern Tungus. This requires further research.

[122] A parallel development **a* > *ï* can be found in the neighbouring Turkic language Salar.

BaoD	BaoÑ	Kgj	CM	
alə	*alar*	*anla*	**ala*	crotch
aldə	---	*anda*	**alda*	fathom
sarə	*sara*	*sara*	**sara*	moon
martə-	*marta-*	*marta-*	**marta-*	to forget
χuinə	*χi:na*	*ġuaina*	**koïna*	behind
orə-	*ər-*	*uru-*	**ora-*	to enter

This development may even happen to long **aa*, see 3.12.2.

3.2.4. Non-harmonic rounding

Unaccented CM **a* may result in a labial vowel in QG in several ways that are unrelated to harmonic rounding. Some of the processes are discussed below.

Rounding of **a* by preceding consonant

Rounding of **a* > *u* under the influence of a preceding labial consonant is especially frequent in the first syllable in Mongghul, as in *bura:-* < **bara-* 'to finish', *muṣda:-* < **(u)marta-* 'to forget'. It is also found in the other QG languages, e.g. EYu *munaŋ ~ manaŋ* < **manan* 'fog'; MgrM, Dgx and Kgj *puda-* < **bagta-* 'to fit into'; MgrH *pusa*, MgrM *puʒa*, Dgx *pəsə ~ pusə* 'also' < **basa*.[123] There are also occurrences in non-first syllables, e.g. Dgx *awuğa* 'uncle' < **abaga*. In Baoan dialects, where even accented **a* can be reduced to *ə*, this reduction product can be rounded by the preceding consonant, as in BaoGt *yȧmu* < **yamə* < **yama* 'meal'.

Rounding of **a* by absorption of **b*

This is not uncommon in QG. It involves the absorption of intervocalic *-b-* as in Kgj *χər- ~ χʉr-* < **kabïd-* 'to swell', MgrH *to:rġa* (< **tabargan*) < **tarbagan* 'marmot', or of preconsonantal *-b-*, as in MgrH *ɕuġua:-* < **ǰabka-* 'to lose'.

Rounding of **a* by following vowel

Rounding of **a* under the influence of the following syllable is relatively uncommon, and occurs mostly in Baoan when the second syllable vowel has an original contraction length. Examples include BaoD *χutuŋ kuŋ* < **katun (+ *küün)* 'woman', and *ćüχor* < **čakïur* 'lighter'. The same phenomenon can be observed in other Baoan dialects and Kangjia. Similar words in Dongxiang may feature reduction between strong consonants, typically in the shape of *ï*, but lacks the rounding.

[123] Another example, MgrM *muqa* 'meat', probably developed from earlier **maqa* (cf. MgrH *maxa*), rather than directly from CM **mïkan*.

BaoD	BaoÑ	Kgj	Dgx	CM	
χuśüŋ	*ġaśəŋ*	---	*qïʂuŋ*	**gasïun*	sour, bitter
χaluŋ	*χələŋ*	*χulɔ*	*qaluŋ*	**kalaun*	hot
χotoŋ	*hdəŋ*	*χutuŋ/χuduŋ*	*qïduŋ*	**kataun*	hard

Other incidental cases include MgrH *suʑog* < **sačug* 'tassel', EYu *duğu:l-* < **daga.ul-* 'to cause to follow'. Prehistoric cases such as **tamu-* > **tomu-* 'to twist', **namugan* > **nomugan* 'docile', **dolaan* < **doluan* < **daluan* 'seven', feature *o* rather than *u*.

Rounding (unknown causes)

Some labialisations of unnaccented **a* lack an obvious explanation, e.g. EYu *ura:* < **arïa* 'molar', EYu *čüsa-* < **ǰasa-* 'to make'; MgrH *dura:sə* < **darasun* 'alcoholic drink', Dgx *arunʑa-* < **aralǰï-* 'to trade'.

Others have more than one possible explanation, e.g. in MgrH *maŋġusə* ~ *moŋġusə* 'ogress' the latter variant could either be caused by the *m-* or by the following vowel (CM **maŋgus*). Both explanations may also apply to Dgx *boruŋ* < **baraun* 'right hand side' and *pudu* < **batu* 'strong'. CM **a* in the middle syllable can easily be assimilated to the preceding or following vowel, as in Dgx *pudura-* < **butara-* 'to be scattered'.

The rounding of **a* followed by *-ŋ* seen in some Baoan dialects and in Kangjia may be related to the Tibetan inspired simplification of vowel diversity preceding *ŋ* (see 3.3.5. and 4.8.3.). Examples:

BaoÑ	BaoGt	Kgj	Dgx	CM	
aćaŋ	*aćɔŋ*	*aćɔ*	*aɕaŋ*	**ačïan*	load
altaŋ	*altɔŋ*	*antɔ*	*antaŋ*	**altan*	gold
bayaŋ	*bayɔŋ*	*bayɔ*	*bayaŋ*	**bayan*	rich
ćiχaŋ	*ćixɔŋ*	*čiχɔ ~ čɔχɔ*	*ɕïğaŋ*	**čagaan*	white

3.2.5. Palatalisation phenomena

CM simple **a* can be palatalised by adjacent sounds in several ways. In the QG languages most cases of palatalisation are due to preceding palatal consonants.

Umlauts and glides

The partial palatalisation of **a* of the first syllable due to /i/ < **ï* in the following syllable occurs in central Mongolic as well as in Dagur.[124] In Dagur, /*a*/ of the first

[124] In most dialects of Mongolian proper and Buriat /*a*/ < **a* is slightly palatalized, which is not indicated in the orthography. However, in several dialects including Chakhar this has led to the creation of additional vowel phonemes with front pronunciations such as *ɛ*. This did not disturb vowel harmony, as the vowel harmony system of these dialects is not of

syllable is also slightly palatalised before /i/ of the following syllable, which will not be indicated here. It does not affect vowel harmony. This type of umlaut is only rarely encountered in the QG languages, as in MgrM *tebi-* < **talbï-* 'to put', Kgj *ǰeǰi-* < **ǰaǰïl-* 'to chew'.

CM **a* may develop a palatal coarticulation after a palatal consonant. This takes the shape of a palatal glide between the consonant and the following vowel (not written by most authors). In Dahejia Baoan a glide *i* or *ĭ* is usually inserted between the alveolo-palatal affricates *ć* and *ʒ́* and the following /a/:

BaoD	CM	
ćiadə-	**čad-*	to be satiated
ćiasuŋ	**časun*	snow
ʒ́iaʒ́ial-	**ǰaǰïl-*	to chew
ʒ́ialχə-	**ǰalkï-*	to swallow

This glide is also described for Mongghul. Mongghul also features a partial palatalisation of the **a* itself. Mostaert & de Smedt write the glide + vowel as *iä*; Khasbaatar and Junast (who do not write the glide before *a*) use the notation *æ*. Cf. the development of **čad-*.

Palatalisation by adjacent consonants

The consonants that are able to palatalise the following vowel are **č*, **ǰ*, and **y*. The Dagur developments differ from those in QG, partly due to the different accent placement.

CM **a* of the first syllable in Dagur is not simply palatalised by an initial **č*, **ǰ*, or **y*. The vowel remains unchanged in words like *čad-* < **čad-* 'to be satiated', **ǰar-* < **ǰaru-* 'to use', *yamər* < **yamar* 'what kind of'. However, in CM word stems with two or more palatal elements, **a* tends to result in Dag *e*, sometimes even *i*, e.g. *čerč-* ~ *čirč-* < **čabčï-* 'to chop', *činč* < **čamča* 'shirt', *ǰebʸǰ* < **ǰabaǰï* 'corner of the mouth', *ǰeŋgʸe:* < **ǰaŋgïa* 'knot', *ǰerʸe:* < **ǰarïa* 'hedgehog', *ǰinč-* < **ǰančï-* 'to crush'. The palatalisation of the vowel in Dag *čeɣʸ* < **čag* 'time' also seems to be due to the presence of two palatal elements, alhough the final vowel was not originally there.[125]

In Dagur unaccented **a* in non-first positions tends to develop into *i* under the influence of a preceding palatal **č*, **ǰ*, or **y*, as in *baǰir* < **buǰar* 'dirty', *gaǰir* < **gaǰar* 'earth', *kačir* < **kačar* 'cheek'. In Shirongol we normally find this development only in Baoan, probably due to the fact that Baoan is the only language with

---

the palatal-velar type. In Oirat, palatalisations caused many back-vocalic words to be transferred to the front-vocalic vowel class. The latter applies even to those Inner-Mongolian dialects which developed a real front vowel *ɛ*.

[125] The Dagur shape *čeɣʸ* need not go back to a disyllabic form **čagï*, but may owe its palatalised final consonant to the gen.-acc. form *ćeɣi:*. The same is seen in other stems, including *galʸ* < **gal* 'fire', *garʸ* < **gar* 'hand', but as these words lack further palatal elements, their vowel did not become *e*.

the tendency **a* > *ə* even in the accented final syllable. This leads to forms like BaoÑ *ġaʑar ~ ġaʑir*, BaoD *ġaʑir ~ ġaćir* < **gajar* ‘land’.

In Shirongol and Eastern Yugur the **a* of unaccented syllables may develop into *i* under the influence of a preceding palatal **č, *ǰ,* or **y*.[126] The effect of **y* is also more widespread in that it may also palatalise a preceding **a*, and **a* in non-first syllables. Examples include EYu *ǰiya* < **ǰayaan* ‘fate’, MgrH *ʑiġa ~ ʑæġa* < **ǰaka* ‘collar’, *ćiġa:n* < **čagaan* ‘white’, Kgj *ǰiğasun ~ ǰağasun* < **ǰagasun* ‘fish’. The palatalising **y* may itself disappear, as in EYu *(y)ida-*, MgrH *ida-* ‘to be unable’ < *yida-* < CM **yada-*, EYu (*y*)*ima* < **yama* ‘thing’, and in non-first syllable: EYu *aiğa,* Dgx (*y*)*iğa* (< **ayiğa*) < CM **ayaga* ‘bowl’.

Palatalised **a* > *i* may become part of a new diphthong, as in *biar* < *bəyar* < **bayar* ‘joy’, *tiaġ* < **tayag* ‘cane’. A similar ‘absorption’ of the unaccented vowel by the *-y-* is found in Dagur, e.g. *bais-* < **bayas-* ‘to rejoice’, *k*^{w}*air* < **kuyag* ‘armour’. Due to the different stress patterns the results often look different:

Dagur	EYu	CM	
aiɣə	*aiğa*	**ayaga*	cup
bayin	*bəyan*	**bayan*	rich
nɔyin	*niyɔn*	**noyan*	official

The palalalisation of first syllable **a* is surprisingly also seen in Dagur, where the affected **a* carries the accent, e.g. *čiɣa:n* ‘white’, *ǰila:* ‘tassel’, and *ǰiya:* ‘fate’ from CM **čagaan*, **ǰalaa*, and **ǰayaa*.

This palatalisation of **a* leads to difficulties in reconstruction, in that it may become impossible to determine whether palatal breaking took place. Forms like BaoÑ *yimaŋ* (CM **ïmaan*, Kh *yamaa*) do not prove that palatal breaking did *not* occur, since an older Baoan form may have been **yamaŋ*.[127]

It is not clear whether Dgx *ɕïğaŋ* ‘white’ developed through **čiğaŋ* as well, or merely represents a reduction **čəğaŋ*. The development **a* > *ï* is quite common in Dongxiang, but is mostly seen preceding original strong consonants, and associated with devoicing. Examples include *qïʑa* < **kačar* ‘cheek’, *ɕïʑï-* < **čabčï-* ‘to chop’.

Dagur also features ‘palatal creep’, a more complicated variation on the ‘umlaut’ theme that is also known from central Mongolic. It involves the palatalisation of the /*a*/ of the second syllable of trisyllabic stems by the the original final syllable /*i*/ < **ï*. Examples include Dag *adil*^{y} < **adalï* ‘similar’, *ariɣ*^{y} < **arakï* ‘liquor’. This leftward movement of the palatality has further led to the palatalisation of the first syllable, as in *ǰelbir-* < **ǰalbarï-* ‘to beg’. The *-e-* of *ǰelbir-* would not have been palatalised by the initial **ǰ-* alone, had the vowel *i* < **ï* remained in its original position.

[126] Unlike in Dagur the development **a* > *e* is rare in the QG languages, except when *a* is part of the diphthong **aï*.

[127] See Nugteren forthcoming.

3.2.6. Elision

Elision of **a* via reduction

Loss of **a* preceding a weak consonant may have developed via an intermediate reduction vowel **ə*. However, a variant with the *a-* preserved is often also attested alongside the elision form, as in MgrH *aləma ~ ləma* 'fruit', whereas forms like **ələma* are generally not documented.

Both the loss of initial **a* and **a* between two consonants can lead to initial consonant clusters. Examples:

E. Yugur	*ndağar*	**andagar*	oath
	χwa:r	**kabar*	nose
	čğa:n ~ čəğa:n ~ čağa:n	**čagaan*	white
	ra:ljə-	**aralǰï-*	to exchange
Mongghul	*ləma*	**alïma*	fruit
	ŋġai-	**aŋgaï-*	to open
	dalə	**adalï*	resembling
	dəla-	**aduula-*	to herd livestock
	ra:l	**aral*	island; axle
	ra:	**arïa*	molar
	yaġa	**ayaga*	bowl
Kangjia	*dərasun ~ drasun*	**darasun*	liquor
	ima	**aïmag*	village
Dongxiang	*(y)iğa*	**ayaga*	bowl
	da-	**yada-*	to be unable

Elision of **a* via devoicing

In case of initial vowels that are lost, Eastern Yugur may display *h-*, apparently a relic of a vowel devoiced due to the following voiceless consonant or cluster. Similarly, Mongghul often has an unexpected initial *s-*, *ś-* or *ʂ-*.

E. Yugur	*hsar*	**asar*	village
	hsəra- ~ sra-	**asara-*	to bring up
	hča:n	**ačïan*	load[128]
	hrča ~ arča	**arča*	cypress
	sqa-	**aska-*	to sprinkle

In other cases the initial consonant is dissimilated or lost:

[128] But cf. EYu *ačə-* < **ačï-* 'to load'.

E. Yugur	*hġa-*	**kaka-*	to suffocate
	hta-	**tata-*	to pull
	hsa- ~ sa-	**ǰasa-*	to make
	χǰa:r	**gagčaar*	alone
Mongghul	*χġai*	**gakaï*	pig
	sġal	**sakal*	beard
	sgə-	**sakï-*	to wait[129]
	saġa- ~ sʒaġa-	**hasag(u)-*	to ask
	sġa-	**aska-*	to sprinkle
	śʑa:	**ačïan*	load

In Dongxiang **a* was only lost completely in a few words:

Dongxiang	*sʑï-*	**sačʉ-*	to sprinkle
	cïra- ~ sïra-	**tasura-*	to break
	sïma	**tasma*	thong
	sda-	**tata-*	to pull

A number of words feature similar developments (reduction, elision, preaspiration) in several languages, e.g. **gakaï* 'pig': EYu *ġaɣqai*, MgrH *xġai:*, MgrM *qġai*, BaoD *ġai*, Dgx *qïġəi ~ qïxġəi*. Cf. also **gagča* 'alone' and its derivations, *ǰasa-* 'to make', **takïa* 'chicken', **tata-* 'to pull'. Nevertheless it is usually impossible to assume that the vowel was already lost in the Proto Shirongol period.

Mittelsilbenschwund in trisyllables can be found in all QG languages, as in EYu *halğan ~ halağan*, MgrH *xalġa*, Dgx *haŋġa* < **halagan* 'palm of the hand'. Unlike other elision phenomena, this is often confirmed in the same set of words by all Shirongol languages including Dongxiang. In that set of words the elision can probably be assumed for Proto Shirongol.

3.3. Development of simple **e*

3.3.1. Default development

In Dagur, **e* is represented by *ə*. In the QG languages the normal development in accented syllables is *e* (in Shirongol with the usual allophones *ie*, *ia*, *iä*, etc). Generally **e* is best preserved when accented and in contexts with neutral consonants, as in **dere* 'pillow', **gerel* 'light', **kelen* 'tongue', **ere* 'man', **nere* 'name'.

In Dahejia Baoan the development of **e* in non-first syllables is often *ə* rather than *e*. A similar neutralization is also seen in the Dahejia Baoan development of other vowels in this position. Peculiarly, some of these words behave similarly in Dongxiang, even when other Baoan dialects feature *e*.

[129] Eastern Yugur and Ñantoq Baoan both feature lengthened *a:* in **sakï-*.

Dgx	BaoD	BaoÑ	CM	
erə	*erə*	*erə*	**ere*	man
etə	*etə*	*erte*	**erte*	early
furə	*furə*	*fure*	**hüre*	seed
ewə	*uer*	*ewer*	**eber*	horn
---	*helgə*	*helge*	**heligen*	liver
piśie	*se*	*se*	**büse*	belt
sugie	*suge ~ ge*	*ʂge*	**süke*	axe

The development of **e* to *a* in the first syllable is frequently observed in Dahejia Baoan, without obvious reason, e.g. BaoD *gaʑiəsuŋ* < **gedesün* 'intestines', *kalə-* < **kele-* 'to speak', *kaməl-* < **kemle-* 'to gnaw', *narə* < **nere* 'name', *nda-* < **ide-* 'to eat'. Ñantoq Baoan typically has *e* in these words (see below for **-en* > BaoD *-aŋ*). Mongghul and Mangghuer have incidental cases of *a* < **e*, e.g. MgrH *sarən* < **seriün* 'cool', *χaŋgu-* < **emkü-* 'to hold in the mouth'. Dongxiang also has a number of instances, e.g. *banɕə* < **belčier* 'pasture'.

3.3.2. Harmonic rounding

After **ö* of the first syllable, **e* is rounded. In Eastern Yugur this situation is normally only preserved when the vowel of the first syllable remains non-high. In the Hongshiwo dialect, on which the most elaborate description of Eastern Yugur was based, the **ö* has usually been raised, and the consecutive syllables remain unrounded.[130]

MgrH	Qinglong EYu	Hongshiwo EYu	CM	
gudoli-	*gödöl-*	*gʉdel-*	**ködel-*	to move
ʒo:lon	*jö:len*	*jü:len*	**jöelen*	soft
bo:ro	*pö:rö*	*pü:re*	**böere*	kidney
ʂdogo:n	*hötgön*	*hʉtgwen*	**ödken*	dense
konorʒə	*kölösən*	*kʉlesən*	**köle(r)sün*	sweat

In Dongxiang we find that some words show harmonic rounding, while others do not, without an obvious system, e.g. *kugie* < **köke* 'blue', *gogo-* < **köke-* 'to breastfeed', and *gogo* < **köken* 'breast', with the same phonetic environment. Cf. also *boro* < **böere* 'kidneys', *komoru-* < **kömeri-* 'to overturn', but *goʑielu-* < **ködel-* 'to move', *koliesuŋ* < **kölesün* 'sweat'.

In Kangjia, harmonic rounding is normal, in many cases leading to a vowel sequence *ʉ-ʉ* unlike elsewhere in Shirongol, e.g. Kgj *kʉkʉ* < **köke* 'blue', *nʉχgʉ* < **nöken* 'cave', *sʉgʉ-* < **sögee-* 'to scold'.

[130] This is not uniquely Eastern Yugur. In Khalkha there is a similar restriction, so that *ə* cannot follow *o* (< **ö*), and *o* cannot follow *u* (< **ü*).

In Mongghul there are also cases of harmonic rounding after **ü* of the first syllable, not supported by other languages, e.g. *(u)go* < **üge* 'word', *sgo* < **süke* 'axe'. However, in some stems this is more widespread, probably due to lowering of the original **ü* to **ö* near velars. This may be the case in EYu *nökön*, MgrH *noko*, MgrM *nuko*, BaoD *nokuŋ*, Kgj *nʉχgʉŋ*, all ultimately from **nüken* 'hole, cave', and in EYu *hgor*, MgrH *fugor* ~ *xgor* , MgrM *xugor*, BaoD *fgor* ~ *gor*, BaoÑ *өkөr* ~ *өχgөr*, BaoX *xgur*, Kgj *gөr*, all from **hüker* 'bovine'. Dongxiang deviates again with its unrounded second vowel in *nokieŋ* 'pit', *fugie* 'bovine'.

Dagur forms such as *kuk*^{w} < **köke* 'blue', *xukur* < **hüker* 'bovine' are not due to harmonic rounding but to a rule that causes all vowels except **i* to become *u* after a rounded vowel.

3.3.3. Reduction/raising

Unaccented *e* may be reduced to *ə*. In Eastern Yugur this neutralisation can be seen occasionally, as in *kərəi* < **kerie* 'crow', Qinglong dialect *əlǰiɣen* < **elǰigen* 'donkey'. It is common in Mongghul, as in *kəle* < **kele* < **kelen* 'tongue'. Often the reduced vowel is raised and fronted to *i*, especially in Narin Guol. Mongghul examples include *imel* < **emeel* 'saddle', *ire* < **ere* 'man'. Kangjia examples include *išǰɔ* < **ebčiün* 'chest', *ire-* < **eri-* 'to look for', *iǰɔ* < **eǰen* 'master'. Some words show this development in several languages, e.g. EYu *kəseg*, MgrH *kəʒaġ* , Dgx *kiśie*, all from < **keseg* 'piece'.

Raising and fronting occurs in **yesün* < **yersün* 'nine' in many languages[131], including both Dagur and most Shirongol dialects, due to the effect of the *y-*. Simultaneously the *-s-* caused devoicing phenomena in several QG languages. Forms: Dag *is*, *yis*, EYu *ǵisən*, *šisən*, MgrH *ʂʒən*, BaoÑ *yirsөŋ*, BaoX *isoŋ*. Similar raising and devoicing effects can be seen among the very erratic reflexes of CM **yeke* 'big'.

3.3.4. Non-harmonic rounding

CM **e* may result in a labial vowel in QG in several ways that are unrelated to harmonic rounding discussed above. These types of labialisation usually occur in the unaccented non-last syllables.

Rounding of **e* by preceding consonant

In parallel with the development **a* > *u*, QG languages also feature rounding and raising of **e* > *u*. This phenomenon is especially frequent in the first syllable in Mongghul, e.g. *bulen* < **belen* 'ready', *mude-* < **mede-* 'to know', *muŋgi* < **meŋge* 'naevus', *mućin* < **bečin* 'monkey', but may also occur in the remaining syllables, e.g. *imu* < **eme* 'woman', *mugen* < **emegen* 'old woman'. As usual, variants without the labialisation often persist as well. The development of EYu *me:ndə* ~

[131] Kalmuck also has *yisn*, whereas Mongolian proper *yösön* and Buriat *yühen* feature rounding under the influence of the second syllable.

mö:ndə < **mendü* ‘peace’ can probably also be attributed to the initial consonant, although the previously rounded second vowel may have played a role.[132]

Rounding of **e* by absorption of **b*

This development involves the disappearance of intervocalic *-b-*, which may then leave the vowel rounded, as in MgrH *to:rə ~ te:rə-* < **teberi-* ‘to embrace’, or of preconsonantal *-b-*, as in MgrH *tudie:- ~ təbde:-* < **debte-* ‘to soak’. There are often several variants in evidence, as in MgrH *udəg ~ vudəg ~ idəg* ‘knee’ < **ebüdüg*. Similar developments can be found in Dongxiang, e.g. *odəu* ‘knee’ < **ebüdüg* , *osuŋ* ‘grass’ < **ebesün*.[133]

Rounding of **e* by following vowel

Rounding of **e* under the influence of the following syllable is relatively uncommon in QG. It is most frequent in Eastern Yugur, typically in words with a rounded contraction length in the second syllable, e.g. *dölü:n* < **deliün* ‘spleen’, *orʉin*, *orü:n* < **ereün* ‘chin’, *sʉrʉin*, *surü:n* < **seriün* ‘cool’. Eastern Yugur causatives may develop forms that deviate from the base verb, e.g. *hergə-* < **hergi-* ‘to turn’ has the causative *xorgü:l-* < *hergi.ül-*.[134]

Cases from Shirongol languages include MgrM *kugor*, BaoÑ *kurgur*, Kgj *kʉrgi ~ kʉrki*, all < **geškiür* ‘stairs’[135], and MgrH *turu:*, BaoÑ *tərəŋ*, Kgj *tʉrʉ*, Dgx *ćiauruŋ*, all < **teriün* ‘head’. The Dongxiang form seems to derive from a metathesised form **teurun*.[136] In case of Kgj *mʉtʉ*, Dgx *mutu* < **metü* ‘like’, the *m-* may have played a role as well.

Older instances of rounding by a following vowel include **elgü-* ‘to hang’, **ergü-* ‘to lift’. In this group the rounding is almost universal, modern forms preserving the *e* are rare. Dagur is especially conservative in this regard, e.g. *ərwə:-* as opposed to EYu *oryo-*, MgrH *(u)rgu-*, MgrM *argu-*, BaoGt *urgə-*, Dgx *uğu-*, all from **ergü-* ‘to lift’. Many other similarly structured stems, such as **emüne* ‘front’, **ebür* ‘bosom’, contain labial consonants as well, i.e., two possible sources of labialisation. These words generally have rounded the initial **e* in central Mongolic, but both in Dagur and in QG there are unrounded forms left, e.g. Dag *əməl*, BaoÑ *eməla*, Kgj *emele ~ emle.* CM **temür* ‘iron’, also rounded in central Mongolic, retains the unrounded **e* throughout QG: EYu *temər*, MgrH *təmur*, MgrM *tərmər*, BaoD *təmər*, Kgj *čimə*, Dgx *ćiemu*.

---

[132] The reverse development, unrounding of **ö* near labials is discussed below.

[133] As initial */o/* is pronounced [*uo*] in Dongxiang, one could argue that the **b* is ‘still present’. However, primary **o* and **ö* result in the same pronunciation.

[134] EYu *eje-* < **üje-* ‘to see’ has the causative *ojü:l-* < **üje.ül-*, and *hele-* < **hüle-* ‘to remain’ has the causative *holü:l-* < **hüle.ül-*, but in these cases the causatives may simply preserve the rounded vowel that was unrounded in the base verb.

[135] Due to the rounded second syllable **geškiür* developed differently from its base **geški-* ‘to step on’.

[136] Cf. also Dongxiang *ćiauru-* < **teberi-* ‘to embrace’ and *ćiauŋgu-* < **temgü-* ‘to pick up’.

In case of CM **üdür* (*?*edür*) ‘day’ and **übül* (*?*ebül*) ‘winter’ the evidence for an original form with **e-* is exceptionally weak, as both Middle Mongol and Dagur suggest a labial first vowel.

3.3.5. Tibetan-inspired simplification preceding *-ŋ*

In Baoan not all vowels are allowed to precede *-ŋ*. This seems to be inspired by Amdo Tibetan, where similar restrictions exist. As Baoan *-ŋ* also represents the very frequent CM **-n*, a large number of words is affected. We get *-aŋ* instead of expected *-eŋ*, and *-uŋ* instead of expected *-oŋ* (with harmonic rounding). In Dahejia Baoan, the development **e* > *a* also tends to occur in the first syllable of these words. In view of other Baoan dialects this is perhaps a recent assimilation phenomenon. Examples:

BaoD	BaoÑ	CM	
kalaŋ	*kelaŋ*	**kelen*	tongue
tarmaŋ	*termaŋ*	**teermen*	mill
balaŋ	*bəlaŋ*	**belen*	ready
kitaŋ	*kitaŋ*	**köiten*	cold
ndaŋ	*ndaŋ*	**eüden*	door
uniaŋ	*unaŋ*	**ünien*	cow
urgaŋ	*werġaŋ*	**berigen*	sister-in-law
kurgaŋ	*kurġaŋ*	**kürgen*	brother-in-law
moruŋ	---	**mören*	river
nokuŋ	*nəkəŋ*	**nöken*	hole

In Kangjia word-final CM **-en*, like **-an*, has become *-ɔ* (via *-ɔ̃*), as in *kiliɔ* < **kelen* ‘tongue’, *ʉniɔ* < **ünien* ‘cow’, *idɔ* < **eüden* ‘door’, *bergɔ* < **bergen* ‘sister-in-law’, *mʉrɔ* < **mören* ‘river’.

3.3.6. Elision

Elision of **e* via reduction

Loss of **e* preceding a weak consonant may have developed via an intermediate reduction vowel **ə*. In the first vowel position, both the loss of initial vowels and those between two consonants can lead to initial consonant clusters. The loss of the vowel before *w* is common in Eastern Yugur. Examples:

E. Yugur	*βesən*	**ebesün*	grass
	βədəg	**ebüdüg*	knee
	(ə)ɣečə	**egeči*	elder sister
Mongghul	*ŋge*	**eŋger*	flap of garment
	mugen	**emegen*	old woman

Mongghul	*nde:*	**ende*	here
	ʒige	**eljigen*	donkey
Kangjia	*vəide-*	**ebed-*	to be ill
	vəidəʉ	**ebüdüg*	knee
	vəisʉn	**ebesün*	grass

The following words lost the initial **e* in several QG languages. Ñantoq Baoan deviates here, in that it sometimes preserved the initial **e* in words in which even the normally conservative Dongxiang lost it. This makes it impossible to postulate a Proto QG form with the elision.

**emüdün* 'trousers': EYu *modən*, *mudən*, MgrM *modo*, BaoD *məduŋ*, BaoÑ *emdəŋ*, Dgx *məduŋ*.

**emüs-* 'to put on, to wear (clothes)': EYu *məs-*, MgrH *musə-*, *mosə-*. MgrM *musï-*. BaoD *musï-* , BaoÑ *məʂ-*, Kgj *məsï-*, Dgx *misï-*.

**emüne* 'front': EYu *ölmö*, *ömle*, BaoD *məilə*, BaoÑ *eməla*, BaoX *mələ*, Kgj *emele ~ emle*, Dgx *mələie*. (MgrH *muśi* and MgrM *meʂï*, apparently represent the directive form **emüne-(g)si* 'forward, to the front').

Elision of **e* via devoicing

In Eastern Yugur the modern form may display *h-*, apparently a relic of a vowel devoiced due to the following voiceless consonant or cluster. Similarly, Mongghul often has an unexpected initial *s-*, *ś-* or *ʂ-*. As can be seen, the presence of original **h-* does not make a difference to the modern forms.

E. Yugur	*hče-*	**heče-*	to be emaciated
	hke	**eke*	mother
	čɣe	**ečige*	father
	šə	**hesi*	handle
	škən	**hekin*	head
	š(i)ke	**yeke*	big
Mongghul	*si:*	**ese*	neg. particle
	sargu	**(h)esergü*	opposite
	ʂde	**erte*	early
	sgəl	**sedkil*	heart

The verb **eči-* 'to go' lost its initial vowel in most Shirongol languages: MgrH *śʑə- ~ śə-*, MgrM *śi-*, BaoD *ći-*, BaoGt *śi-*, BaoX *xǰə-*, Kgj *ǰi-*. However, the fact that

Dongxiang retained the disyllabic form *əɕï-* makes it impossible to reconstruct the shorter form for Proto Shirongol.[137] See also the negation particle **ese*.

As in the case of other vowels, medial **e* in trisyllabic stems is often elided in all of Shirongol, as in MgrH *kurge:n*, BaoGt *kurgɔŋ*, Dgx *quğoŋ* < **küregen* 'son-in-law'. In such unanimous cases a disyllabic form may be assumed for Proto Shirongol. In this and other cases Eastern Yugur also suggests a disyllabic form.

3.4. Development of simple **ï*

3.4.1. Preliminary remarks on CM **ï* and **i*

The reconstructions used here are based on an eight vowel system which includes**ï* as an harmonic counterpart to **i*. Evidence from Middle Mongol, Moghol, and the QG languages shows that such a system existed at least in some periods. As elsewhere in Mongolic **ï* and **i* themselves have become mostly indistinguishable in the peripheral languages. They merged into *i* in Dagur, and into *ə ~ i* in the QG languages. The original distinction has left indirect traces in Eastern Yugur and Baoanic, especially in the treatment of **k* and **g*. These consonants have split into velar *k* and *g* and their uvular counterparts *q* and *ġ* near the merged vowel, so that diachronic information about **ï* can be gleaned from these languages. E.g. the reconstruction **čïkï-* 'to stuff into', as suggested by the *-q-* in the Muqaddimat al-Adab and Moghol, is confirmed by EYu *čəqə-*. The development of these consonants suggests that **ï* was a real back vowel, at least in an earlier stage of these languages. However, in other periods or languages **ï* must have been a palatal vowel, as in the Monguor languages both **ki-* and **kï-* sequences resulted in *ć*i-.

In words that do not contain **k* and **g* or other back vowels, it may be difficult to separate **ï* from **i* without circumstantial evidence, such as inflected or derived forms, or non-Mongolic data. The pronouns **bï* 'I', **čï* 'you' were back-vocalic in view of their inflected forms, e.g. acc. **čï.ma.ï* 'you', dat. **čï.ma.da* 'to you'. CM**ǰïl* 'year' may have been back-vocalic **ǰïl* in view of its Turkic cognate **yïl*. The verb **ki-* 'to do' may stem from **kï-*, but the evidence for this is weak.

3.4.2. 'Breakable' and 'unbreakable' **ï* and **i*

The development of **ï*/**i* in first syllables strongly depends on the second vowel in the word. The most straightforward development can be found in **ï*/**i* of monosyllabic stems, **ï*/**i* of non-first syllables, as well as first-syllable **ï*/**i* followed by another **ï*/**i*.

When the second vowel was an **a*, **o*, **ö*, **u*, or **ü*, the **ï*/**i* of the first syllable may undergo so-called palatal breaking, whereby the second vowel is imported into the first syllable, either joining or replacing the original **ï*/**i*.

[137] Poppe (1955:30) assumes this verb derives from **od-* 'to go', in which case it would be an example for an irregular development (and elision) of **o*.

**i* followed by *e* in the next syllable does not undergo breaking in central Mongolic, but in Dagur, Eastern Yugur, and Mongghul, words with this vowel pattern occasionally do undergo assimilations that are superficially reminiscent of prebreaking (see 3.6.1.).

In the following we will first deal with 'unbreakable' **ï*, then with 'unbreakable' **i*, and finally with breaking of **i*/**i* in back and front-vocalic stems.

3.4.3. Default development of **ï*

'Unbreakable' **ï* and 'unbreakable **i* have merged into *i* in the three central literary languages (although not in all central Mongolic dialects) and in Dagur. In the QG languages **ï* and **i* , and in many cases also **u* and **ü*, have the tendency to merge into *ə (~ i)* in non-first syllables. This tendency is shared by other languages of the *Sprachbund*, including Amdo Tibetan and the Turkic languages Salar and Western Yugur. This tendency can be called 'high vowel neutralisation'. Examples for the development of **ï* in monosyllables and following various back vowels: EYu *čə*, MgrH *ćə*, MgrM *ći*, BaoÑ *ći*, Kgj *či* < **čï* 'you (sing.)', EYu *amən*, MgrH *amən* ~ *amun* < **amïn* 'life', *narən* < **narïn* 'fine', EYu *ġučən* ~ *quǰin*, MgrH *xoǵin* < **gučïn* 'thirty', EYu *šɔrğɔlǰən*, MgrH *śirġɔlǯin* < **sïrgolǰïn* 'ant'. As seen here, *ə* may appear as the allophone [*i*] when preceded by a palatal consonant. In Eastern Yugur the palatal pronunciation seems to be optional, cf. *aǰərğa ~ aǰirğa* < **aǰïrga* 'stallion', *ǰaǰəl- ~ ǰaǰil-* < **ǰaǰïl-* 'to chew', *ǰərğa- ~ ǰirğa-* < **ǰïrga-* 'to be glad', but even after palatal consonants *ə* is the more common pronunciation.[138]

In Dongxiang the back vowel *ï* tends to occur not only after uvular *q* and *ğ*, but, in some analyses, also after apicals *s*, *c* and retroflexes *ʂ*, *ʐ*, *ʒ*. In this analysis, which will be followed here, *ï* may be viewed as a separate vowel phoneme, which however should not be seen as a continuation of the old phoneme **ï*.[139] The actual pronunciations following the apical and retroflex consonants are [ɿ] and [ʅ], respectively. In Mangghuer, which like Dongxiang is strongly influenced by Chinese, the same [ɿ] and [ʅ] can be found following apicals and retroflexes, but *ï*-like pronunciations do not occur near the uvular consonants.[140]

3.4.4. Rounding phenomena

Rounding of **ï* by preceding consonant

There are several old alternations **ï ~ *u* near labial consonants for which it is generally hard to determine which is the older variant, such as **kabïd- ~ *kabud-* 'to swell', **kabïrga ~ *kaburga* 'rib'. Isolated modern cases include EYu *muna* < **mïnaa* 'whip', MgrH *xavuʒə* < **kabïsun* 'rib', MgrH *amun* < **amïn* 'life', MgrM

---

138 This does not apply when *i* goes back to long *i:*.

139 In an alternative Chinese-inspired analysis [ɿ] and [ʅ] are attributed to the phoneme /i/ and written <i>.

140 In Baoan dialects with strong Chinese influence and in Kangjia, *ï* occurs following apicals and/or preceding gutturals. However, the modern *ï* in these languages rarely goes back to CM **ï*.

amula < **alïma* ‘fruit’. In several languages: EYu *buda*, MgrH *buda*, MgrM *buda-ŋ* < **bïda* ‘we’.

MgrM *muqa* and BaoX *moġa* ‘meat’ seem to stem from a form **maqa* with prebreaking, rather than directly from CM **mïkan*, in view of the cognates with *-a-* in Mongghul and Baoan dialects.

**ï* > *u* preceding *-ŋ* in Baoan and Dongxiang

In Baoan, the sequence *-iŋ* (the expected reflex of **-ïn*/*-in*) simply seems to have become inadmissible, perhaps due to Tibetan dialects which show similar restrictions preceding *-ŋ*. However, such an explanation would not apply to Dongxiang, which is not otherwise known to have undergone Tibetan influence. The **ï* is either lowered > *a* or rounded > *u*. Interestingly, Kangjia appears to agree mostly with the Ñantoq Baoan forms, given that Kangjia *-ɔ* stems from **-an* (in combinations *-un*/*-ʉn* the *-n* is not lost in Kangjia).[141]

BaoD	BaoÑ	Kgj	Dgx	CM	
amuŋ	*amoŋ* (BaoX)	*amin*	*amiŋ*	**amïn*	life
labćüŋ	*labćaŋ*	*lašǰɔ*	*laçəŋ*	**nabčïn*	leaf
naruŋ	*na:raŋ*	*narɔ*	*naruŋ*	**narïn*	thin
ćiχaŋ	*ćiχaŋ*	*čiχɔ*	*çiġəŋ*	**čïkïn*	ear
ʒ́ioćüŋ	---	*ǰʉčɔ*	*ʒoçəŋ*	**ǰočïn*	guest
χuićüŋ	*χi:ćaŋ*	*χuaičɔ*	*quaiçəŋ*	**kaučïn*	old
χoroŋ	*χərəŋ*	---	*qoruŋ*	**korïn*	twenty

If the **-n* is lost, **ï* does not result in Baoan *u*, e.g. BaoD *ġonə*, BaoÑ *ġəni* < **konïn* ‘sheep’, BaoD *morə* < **morïn* ‘horse’, and likewise if there was no final consonant, e.g. BaoD *ʒ́ialχə-* < **ǰalkï-* (< **ǰalgï-*) ‘to swallow’, *ćiχə-* < **čïkï-* ‘to stuff’. This is not surprsing, given that even original **u* is unrounded if **-n* is lost, e.g. BaoD *nasə* ‘age’, as opposed to Dgx *nasuŋ* < **nasun*). There are cases where this vowel change occurs in Dongxiang with other final consonants:

BaoD	BaoÑ	Kgj	Dgx	CM	
---	*yarəm*	*yarun*	*ʒaruŋ*	**ǰarïm*	half
---	---	---	*ġuruŋ*	**kurïm*	feast
ġulər	*ġələr*	*gʉrʉ*	*ġuruŋ*	**gurïl*	flour

The rounded first syllable or the final **m* (rather than original *n/ŋ*) may have played a role in these cases.

[141] Kgj *išǰɔ* ‘chest’ also agrees with BaoÑ *ebćaŋ*, and in this case, with Dgx *əçan,* whereas BaoD *ebćüŋ* retains (or redeveloped) the high vowel of CM **ebčiün*.

Rounding of **ï* due to various causes

Labialisation of first-syllable **ï* by the second syllable vowel constitutes palatal breaking (for which see below). There are some other sources for labialisation.

Occasionally the **ï* is rounded by the vowel of the preceding syllable, as in EYu *dɔgšun* < **dogsïn* 'fierce', *hčʉr* < **učïr* 'occasion', MgrH *śʐüra:-* < **učïra-* 'to meet'.

Rounding may be caused by a following consonant, as in BaoÑ *xɵwar*, Dgx *ʂuwa* < **sïbar* 'mud', MgrM *ʗubar* < **kïlbar* 'easy'.

Some roundings in Dongxiang seem to be due to the connective vowel *u* which was reinterpreted as part of the stem: *arulu-* < **arïl-* 'to clear up'[142], *ʐaʐulu-* < **ǰaǰïl-* 'to chew' (cf. Dgx **çudu-* < **čad-u-* discussed above).

Inexplicable labialisations include EYu *ǰuda ~ ǰəda* < **ǰïda* 'spear', *qutad* < **kïtad* 'Chinese', *χurğa- ~ χərğa- ~ χarğa-* < **kïrga-* 'to shear'[143], and MgrH *ʑül* (normally *ʑil*) < **ǰïl* (*?*ǰil*) 'year'.[144] The equally unexplained labialisation of the 'converter' suffix **-kI* > *-ku* is a shared feature of the Shirongol languages.

3.4.5. Elision

If no other variants are attested, elision makes it impossible to establish whether palatal breaking ever took place. Related languages may have quite different forms. In these cases, as in *člu:* < **čïlaun* 'stone', *hru:r* < **hïroar ~ *hïraur* 'bottom', **ï* may have been broken before being elided. In cases such as EYu *šlə* we can exclude breaking, but then we do not know whether the CM form was **sili* or **sïlï* 'nape'. Whenever palatal breaking would result in a non-high vowel, it is less likely that this would subsequently be lost. Thus, EYu *ma:n* 'goat' is more likely to derive from an unbroken form, perhaps **əma:n*, than from a broken form **yama:n*.

Elision of **ï* via reduction

E. Yugur	*ma:n*	**ïmaan*	goat
	nagšə	**ïnagsï*	hither[145]
Mongghul	*naġʂə*	**ïnagsï*	hither
	ləġa-	**ïlga-*	choose
Dahejia Baoan	*bar ~ ʂbar*	**sïbar*	mud
	naġa	**sïnaga*	ladle

[142] It seems unnecessary to assume influence from the Dongxiang adjective *aruŋ* < **arïun* 'clean'.

[143] Cf. the cases of unexpected *u* < **a* in Eastern Yugur *ura:*, *čüsa-* discussed above.

[144] Perhaps the Mongghul form stems from a derived form, cf. Dgx *ʐuŋtu* < **ǰïl.tu* 'born in the the year ...', whereas *ʐəŋ* < **ǰïl* 'year (of the animal cycle)'.

[145] This stem is one of the rare cases that also may lose the initial vowel in Central Mongolic languages.

Elision of **ï* via devoicing

E. Yugur	*č(ə)na-* (~ *čəna-*)	**čïna-*	to cook
	š(ə)ra	**sïra*	yellow
	čkən ~ čqən	**čïkïn*	ear
Mongghul	*ʂda:-*	**sïta-*	to catch fire
	śʑu:r	**hïǰaur*	root
Ñantoq Baoan	*ɤ́i*	**bïsï*	(is) not

**čïda-* 'to be able' usually loses its first vowel in the QG languages: EYu *šda-*, MgrH *ʂda-*, BaoD *da-*, BaoÑ *ʂda-*, BaoX *da:-*, Kgj *š(i)da-*.

Mittelsilbenschwund is found in longer stems, usually between syllables with high vowels. Examples include EYu *alğa-* < **arïlga-* 'to clean' and Baoan *almaŋ* < **alïma* 'apple'.

Especially after *-r-*, **ï* and other high vowels may disappear in word-final position, as in EYu *uğwa:r* <**ugaa.rï* 'washing water', and formations with the deverbal noun suffix *-bUrI*.[146] Another example is EYu *ǰa:r, ǰa:rə* < **ǰïar(ï)* 'musk (deer)', but in this word the final vowel is not unanimously supported by other languages either. The same may be seen in verb stems, such as EYu *bar-* < **barï-* 'to grab', *χa:r-* 'to geld' < **kaarï-* 'to scorch', *šər-* < **sïrï-* 'to stitch'. Occasionally also in Shirongol, e.g. BaoÑ *tar-* < **tarï-* 'to plant', Kgj *utur-* < **udurï-* 'to lead'.

3.5. Development of simple **i*

3.5.1. Default development of **i*

As mentioned above in 3.4.3. 'unbreakable' **ï* and **i* have merged into *i* in the three central literary languages and Dagur. In the QG languages **ï* and **i* have the tendency to be depalatalized into the vowel *ə*, although the pronunciation *i* may be preserved (or redeveloped), especially when preceded by a palatal consonant. In non-first syllables **u* and **ü* have the tendency to merge into *ə* as well, thus leading to the neutralisation of the high vowels. This is related to the partial loss of vowel variations in suffixes originally containing *U* (**u*/**ü*) or *I* (**ï*/**i*). Although the main reflex of the high vowels is the, phonetically mid, vowel *ə*, the high vowels still form a distinct category, often freely alternating with *i* and *u*, or *Ø*. Examples for the loss of palatality include EYu *pəčə-* < **biči-* 'to write', *əlβə-* < **ilbi-* 'to grope', EYu *setgəl*, MgrH *sgəl* < **sedkil* 'mind', MgrM *ćirəġ*, BaoÑ *ćerəg* < **čerig* 'soldier'. The influence of palatal and labial consonants is demonstrated by the following examples: EYu *belčir* < **belčir* 'confluence', MgrM *puźiġ* < **bičig* 'writing', EYu *döčən ~ döčin*, MgrH *təźin* < **döčin* 'forty'. However, some words widely preserve

[146] The same tendency to drop vowels after *-r-* exists in Western Yugur, with occasional parallels in North Eastern Turkic, including derivations with the suffix *-GArU* (see Nugteren & Roos 2006:118-119).

the pronunciation *i* even without palatal consonants, e.g. BaoD *niŋgaŋ*, Kgj *niŋgɔ*, Dgx *niŋkien* < **nimgen* 'flimsy', and MgrH *ir* < **hir* 'edge of a knife'.

Dongxiang and Mangghuer have *ï* after retroflex consonants, as in Dgx *ʐïrəŋ* < **ǰirim* 'bellyband', *ʂïrə* < **siree* 'table', MgrM *ʂïni* < **sini* 'table'. Words like these also demonstrate there is no connection between this back *ï* pronunciation and CM **ï*.

3.5.2. Rounding phenomena

Most cases of labialisation are caused by labial consonants that either follow or precede the **i*, cf. MgrM *ʂumi-*, MgrH *śimu-*, both from CM **simi-* 'to suck'. In words with rounded vowels 'normal' labialisation is hard to distinguish systematically from palatal breaking (see below). Similar developments occur throughout Mongolic and even within Middle Mongol (and in both directions). The alternants **hüleü ~ *hileü* 'more than' may be explained by labial consonants as well, if the alternation stems from the period before PM **p-* had become **h-*.

In some words the **i* was broken or assimilated in several dialects. For instance BaoX *su:kə*, Kgj *səʉki ~ sikəʉ*, Dgx *sumuğə* 'awl' seem to suggest a development from a broken form **sübüge* (< **sibüge*), which must have developed early enough to prevent the development **s-* > *š-*. However, forms in closely related dialects, such as BaoD *śibġu*, BaoÑ *śuke*, make it impossible to place this development at an early date. Cf. also the Dagur words whose **s-* did not palatalise due to early metathesis of diphthong elements, as in Dag *suidur* < **siüderi* 'dew' (see under **s* in 4.6.1.)

Labialisation can also be caused by the vowel of the preceding syllable, as in EYu *hkon*, MgrH *fuʑün ~ śʑün* < **ökin* 'girl', MgrH *nućün ~ noćin* < **önečin* 'orphan', Kgj *tʉrgʉ-*, Dgx *tuŋku-* < **tülki-* 'to push'.

3.5.3. Lowering of **i*

Lowering of first-slot **i* before **e*

In Eastern Yugur there is a development reminiscent of palatal breaking in words with an **iCe* vowel sequence, which will develop into > *e-e*. Examples include *ere-* < **ire-* 'to come', *eryen* 's/he' < **irgen* 'people', *ede-* < **ide-* 'to eat', *neŋgwen* < **nimgen* 'flimsy', *šere* < **siree* 'table'. The fact that this change also occurred in Turkic loanwords in Eastern Yugur such as *beleg* 'wrist' and *elmeg* 'button loop' (from CT **biläk* and **i:lmäk*) suggests that it is a fairly recent development not connected to palatal breaking.[147] The unaccented first syllable *e* may in fact be a secondary development of *ə*. There are some examples where Qinglong dialect features *ə* instead of Hongshiwo *e*, e.g. Qinglong *nəye* 'one' for Hongshiwo *neye* <

[147] The similarity with Buriat in fig. 21 in Svantesson et al. (2005:196) is coincidental. Unbroken **i* of the first syllable always results in Buriat *e* irrespective of the following vowel.

**nige* ‘one’, Qinglong *šəŋgen* for Hongshiwo *šeŋgen* < **siŋgen* ‘watery’.[148] The same development is found in a number of words in Mongghul: *nengen* < **nimgen* ‘flimsy’, *śæŋgan* < **siŋgen* ‘watery’, *xerge* < **irge* ‘wether’. EYu *her* < **hir* ‘edge of a knife’ is a peculiar case because the *e* was not triggered by the second syllable.

Lowering of **i* preceding *-ŋ* in Baoan and Dongxiang

Parallel to the development of words ending in *-ïn*, words in *-in* undergo a development > Dgx *-əŋ*, Bao *-aŋ*, Kgj *-ɔ*, as in BaoÑ *debśaŋ*, Kgj *dešɔ* < **tübsin* ‘flat, level’, BaoÑ *bećaŋ* < **bečin* ‘monkey’, Dgx *onieçəŋ* < **önečin* ‘orphan’, Kgj *vəićɔ* < **ebečin* ‘illness’. However, the occasional rounding seen in words with *-ïn* is not mirrored here.

3.5.4. Elision

As in the case of ending in **ï*, it is impossible to say whether palatal breaking took place before the vowel was elided.

Elision of **i* via reduction

E. Yugur	*βəl-*	**ibil-*	(of milk) to flow
	šlə	**sili*	nape
Mongghul	*de-*	**ide-*	to eat
	re-	**ire-*	to come
	rgen	**irgen*	people

Elision of **i* via devoicing

E. Yugur	*hni:-*	**hinie-*	to laugh
	hsə-	**is-*	to ferment
	hteɣe-	**itege-*	to believe
	šdən	**sidün*	tooth
Mongghul	*śʑe:-*	**hiče-*	to be ashamed
	śüre:-	**bisire-*	to believe
	sgi:	**isegei*	felt
Kangjia	*ǰasun*	**hičesün*	willow

In non-first syllables **i* may disappear under the same circumstances as the other high vowels. An example for *Mittelsilbenschwund* is MgrM *xarge*, BaoD *helgə* < **heligen* ‘liver’. Loss of word-final **i* can be seen after **r*, as in EYu *ö:r* < **öri* ‘debt’, *teŋger* < **teŋgeri* ‘sky’, and also in verb stems, e.g. EYu *e:r-* < **eri-* ‘to

[148] Cf. Bolčuluu & Jalsan 1988:351. That this *e* developed from **i* through a stage *ə* seems to be supported by Eastern Yugur *eje-* < **üje-* ‘to see’ and *hele-* < **hüle-* ‘to remain’.

seek', *kömör-* < **kömeri-* 'to overturn'. Therefore Eastern Yugur is not helpful in resolving contradictory forms in other languages. For instance in the case of **siber* or **siberi* 'foot sweat', where the Written Mongol spelling has *-i* but the Buriat and Ordos forms end in *-r*, the Eastern Yugur equivalent *šəβer* could stem from either variant.

3.6. Palatal breaking

3.6.1. Assimilation of CM **ï* and **i* to following vowels

Palatal breaking, or *i*-breaking, is the traditional term in Mongolic studies for an assimilation phenomenon whereby CM **ï* or **i* of the first syllable is assimilated to the vowel of the second syllable. This phenomenon occurs when a first syllable **ï* is followed by **a*, **o*, or **u* in the next syllable, or when first syllable **i* is followed by **ö*, or **ü*. This includes lexemes in which **ï* and **i* are followed by double and complex vowels starting with the same elements, such as **aa*, **aï*, **oa*, **üü*, etc. Notably absent from the list of vowels that induce breaking is **e*.

Palatal breaking may manifest itself in two ways. In some stems it is expressed as the replacement of CM **ï* and **i* by the vowel of the second syllable, e.g. Kh *maxan* < **makan* < CM **mïkan* 'meat'. In other stems the second vowel is introduced in the first syllable as well, but the original vowel remains as a 'residual' palatal element, e.g. Kh *myangan* < CM **mïŋgan* 'thousand'.[149] The former type, which in fact entails the complete assimilation of the first syllable vowel, is called 'prebreaking' in Janhunen (2003b:5). This label is applied when there is no residual palatality [in words where it could have been preserved], e.g. Khalkha *maxan* (rather than **myaxan*) < **mïkan* 'meat', Kalmuck *üld* (rather than **yüld*) < **ildü* 'sword'.

Below I will not systematically use the distinction between breaking and prebreaking in the context of the QG languages. As breaking is a phenomenon that affects the vowel of the first syllable, whose phonetic substance and phonemic relevance is generally diminished in the QG languages, it is often difficult to determine whether breaking took place, let alone which type.[150] Some palatal 'residues' may have disappeared only recently, as illustrated by MgrH *narġai* 'tender', as opposed to MgrM *n^{y}arġai* 'id'. Secondary palatalisations may obscure the picture by giving the impression that no palatal breaking took place. Due to its palatal initial consonant, **čagaan* 'white' has resulted in BaoD *ćiġaŋ*, which casts doubt on forms like *śira* < **sïra* 'yellow'. Although this Baoan form looks like a straightforward unbroken continuation of the CM form, it cannot be excluded that it developed via a broken form **śara*. Conversely, secondary labialisations may give the impression that breaking did take place, as in MgrH *buru:* < **bïrau* 'calf'. For

[149] This palatal element is usually analysed as a palatalisation of the initial consonant rather than as a member of a newly-developed diphthong. I am not concerned here with its synchronic phonological status.

[150] Other evidence that would enable us to distinguish breaking and prebreaking is also lacking in the QG languages, e.g. **č* and **ǰ* did not each split into two affricates as they did in the three literary central languages. The word shape of Kh *caray* 'face' shows that **čïraï* 'face' developed into **čaraï* early enough for **č-* to develop into depalatalised *c-*.

these reasons the QG languages are not expected to provide many new insights regarding palatal breaking.

In spite of the abovementioned difficulties, it is clear that palatal breaking is present in a number of words in both Eastern Yugur and Shirongol. The evidence will be presented below, after a look at the more transparent situation in the central languages and Dagur.

3.6.2. Palatal breaking in the North

In the three literary Central Mongolic languages and Dagur, palatal breaking can be most clearly observed, as the affected vowels carry the accent, and are less likely to be assimilated by the consonantal environment. It is worth noting that even Central Mongolic is not monolithic in the application of palatal breaking. Kalmuck retains more unbroken **ï* and **i*'s than Khalkha and Buriat, while Ordos and Khamnigan have very few instances of breaking. Chakhar, which in most respects is quite close to Khalkha, has many cases of unbroken **ï*. Two regular examples in back-vocalic stems:

CM	Kh	Bur	Kalm	Dag	
**sïra*	*šar*	*šara*	*šar*	*šar*	yellow
**čïna-*	*čanax*	*šanaxa*	*čanx*	*šanə-*	to cook

Regular examples with breaking in all these four languages like the above are relatively rare. Incidental cases of breaking are already documented in the earliest documents, and through the centuries developed into a tendency that affected more and more words, apparently without becoming a universal sound shift. During the slow spread of breaking, other sound shifts were taking place, which has led to many inconsistencies. We find that **sïbaun* 'bird' resulted in Khalkha *šuvuu*, but the similarly structured **bïrau* 'calf' resulted in Khalkha *byaruu*, apparently because the contraction of the diphthong **au* was completed at the time of breaking in the first word, but not in the second.

Some irregularities may be explained as the result of mutual influence between the dialects. Other unexpected forms in modern languages may be due to the influence from the written language, which generally retains the spelling with <*i*>.

The examples below from Northern Mongolic clearly demonstrate whether breaking or prebreaking took place in a given stem in a given language. Interestingly, Buriat and Dagur developments often agree. The examples also reveal the internal inconsistencies that can occur within each language, and the resulting irregular correspondences among languages.[151]

[151] That breaking need not be old in Khalkha either is shown by cases like Kh *xyarguy* < **kïrguï* 'hawk', *xyadax* < **kïdu-* 'to annihilate', *nyacax* < **nïču-* 'to retreat', which were apparently only broken after the following *u* was 'neutralised' to *a*.

Breaking of **ï* by **a* of the second syllable

CM	Kh	Bur	Kalm	Dag	
**ǰïda*	*ǰad*	*ǰada*	*ǰid*	[*gʸad* < Ma]	spear
**mïkan*	*max(an)*	*myaxa(n)*	*maxn*	*mʸaɣ*	meat
**mïŋgan*	*myanga(n)*	*myanga(n)*	*miŋɣn*	*mʸaŋgə*	thousand
**nïlka*	*nyalx*	*nyalxa*	*nilx*	*ñalk*	tender
**ïlga-*	*yalgax*	*ilgaxa*	*ilɣăχă*	*yaləɣ-*	to choose
**ïmaan*	*yamaa(n)*	*yamaa(n)*	*yaman*	*ima:*	goat
**hïlaa <*hïlua*	*yalaa(n)*	*ilaahan ~ yalaahan*	*iläsn*	*xila:*	fly

Breaking of **ï* by **o* of the second syllable

CM	Kh	Bur	Kalm	Dag	
**čïno*	*čono(n)*	*šono*	*čon*	---	wolf
**ïro*	*yor*	*yoro*	*yor*	*yɔr*	omen
**ǰïloa*	*ǰoloo(n)*	*žoloo*	*ǰola*	*ǰilɔ:*	reins

Breaking of **i* by **ö* of the second syllable

**čidör*	*čödör*	*šüder*	*čödr*	*šidər*	hobble
**čilöe*	*čölöö(n)*	*sülöö*	*čölän*	*čulə:*	free time
**kiröe*	*xöröö(n)*	*xyüröö*	*körä*	*kirə:*	saw

Breaking of **ï* by **u* of the second syllable

**nïku-*	*nuxax*	*nyuxaxa*	*nuxx*	*nɔɣʷ-*	to knead
**nïsun*	*nus(an)*	*nyuha(n)*	*nusn*	*ñɔs*	snot
**nïruun*	*nuru(n)*	*nyurga(n)*	*nurɣn*	*nirɔ: ~ ñirɔ:*	back

Breaking of **i* by **ü* of the second syllable

CM	Kh	Bur	Kalm	Dag	
**nidü-*	*nüdex*	*nyüdexe*	*nüdx*	*nid-*	to pound
**nidün*	*nüd(en)*	*nyüde(n)*	*nüdn*	*nid*	eye
**sidün*	*šüd(en)*	*šüde(n)*	*šüdn*	*šid*	tooth

3.6.3. Dagur contribution to palatal breaking

As can be seen in the forms listed above, palatal breaking is applied often, but not always by Dagur in back-vocalic stems. Dagur and Buriat share a proclivity to preserve a palatal element more than Khalkha and Kalmuck, cf. Dag *mʸaɣ* < **mïkan*

‘meat’, *n^{y}ombus* < **nïlbusun* ‘tear’. However, unlike in Buriat, breaking in Dagur is rare in front-vocalic stems. Apart from some cases where broken and unbroken variants are attested, as in **ničügün* ‘naked’, **i* mostly remains unbroken preceding **ü/ö* of the following syllable, as in Dag *nid* < **nidün* ‘eye’, *kirə:* < **kiröe* ‘saw’, *šidər* < **čidör* ‘hobble’. On the other hand, Dagur seems to feature some unique cases of breaking before a following **e*, e.g. Dag *šəur* < **siberi* ‘foot sweat’. Some similarly structured Dagur forms do not align with conventional reconstruction. Dag *yəul-* ‘to flow (of milk from the udder)’ may point to a CM form **ibel-* (rather than **ibil-* as suggested by other languages). *šəuɣ^{w}* ‘awl’ may indicate that **sibüge* ‘awl’ is a secondary development of **sibege*.

3.6.4. Eastern Yugur evidence for palatal breaking

As mentioned above, palatal breaking is documented for the QG languages, but it may be difficult to establish whether it has taken place in a given individual word. Even if the vowel of the first syllable has survived, its quality is likely to be influenced by the preceding consonant. In Eastern Yugur and Shirongol the vowel, whether it was broken or not, can be elided or reduced to a degree that makes the word unusable as evidence, e.g. EYu *šdən* < **sidün* ‘tooth’.

But apart from the many words that are indecisive, some clearly demonstrate that breaking did take place in these languages. These are mostly words with a structure that is not conducive to vowel elision, such as words starting with a nasal and with a plosive as the second consonant, e.g. EYu *maɣqan*, MgrH *maxa* from **mïkan* ‘meat’. But it is hard to get unanimous cases; in this very same word the Dongxiang cognate *miğa* stands out with an unbroken *i*.

A number of words clearly suggest that breaking did not take place, such as EYu *čəna* < **čïno* ‘wolf’, instead of expected **čono*.

Among the more straightforward cases of breaking in Eastern Yugur are the following.[152] The word shapes do not necessarily correspond to those in Shirongol.

E. Yugur	CM	
maɣqan	**mïkan*	meat
nagta	**nïgta*	precise
nulusun	**nïlbusun*	tear
nudurğa	**nïdurga*	fist
nuru:n	**nïruun*	back
čusun ~ čüsən	**čïsun*	blood
šɔrui	*?*sïroï* < *sïruaï*	earth
oldə, uldə	**ildü*	sword
nodʉ-	**nidü-*	to knead

[152] It is not clear whether some of these words may have had palatalised consonants at an earlier stage, e.g. *nudurğa* < *ńudurğa* ‘fist’, or *maɣqan* < *ńaɣqan* ‘meat’.

(continued)

E. Yugur	CM	
nʉdʉn	**nidün*	eye
bodü:n	**bidüün*	coarse
šorʉin	**sirüün*	coarse
kʉre:	**kiröe*	saw
hörö:r	**hiröer*	prayer
čüder	**čidör*	tether
čölö:	**čilöe*	spare time

Several words have unexpected forms. EYu *qutağa* < **kïtuga* 'knife', is strange in that it does not have *u* in the second syllable. The lowering of the first vowel of *bodü:n* and *šorʉin* is also unexpected. It may be related to the unexpected *e* in these words in several Shirongol languages, e.g. Dgx *bieduŋ* and *śieruŋ*. EYu *hörö:r* 'prayer' represents a prebroken form as in Ordos *örö:l*, while Khalkha has a broken form and retains a palatal residue in *yörööl*.

In the following cases the Eastern Yugur data are inconclusive. This is either due to reduction or elision of the first syllable vowel, or to the coexistence of variants that contradict one another. (Note that *ə* is the normal Eastern Yugur reflex of both **ï* and **i*, and not infrequently of **u* and **ü*).

E. Yugur	CM	
čna- ~ čəna-	**čïna-*	to cook
šra ~ šəra	**sïra*	yellow
člu: ~ čəlu:	**čïluun*	stone[153]
ma:n	**ïmaan*	goat
šdən	**sidün*	tooth
šu:n	**sïbaun*	bird
šəβɣe ~ šʉɣe	**sibüge*	awl
hru:r	**hïroar/*hïraur*	bottom
χəməsən	**kïmusun*	nail

In some cases the phonetic environment does make it possible to determine whether 'breaking' has taken place. In cases like EYu *məŋgan* < **mïŋgan* 'thousand' and *ġəǰa:r* < **kïǰaar* 'edge', with the vowel between weak consonants, it is less likely for any *a* that might have existed to be reduced.

In cases such as the following **oa* must have become *ɔ:*, and subsequently raised to *u:*. It is impossible to determine whether the variant with *u* in the first syllable is older than the variant with *ə*, or a relatively late development.

[153] Note that this word is usually reconstructed **čïlaun*, but none of the modern languages preserves a trace of the **a*.

E. Yugur	CM	
ǰəlu: ~ ǰulu:	**ǰïloa < *ǰïlua*	rein
ǰirğu:n ~ ǰurğu:n	*?*ǰïrgoan < *ǰïrguan*	six

The following cases are also hard to assess. If they are assumed to be instances of breaking with equivalents in central Mongolic, an Eastern Yugur form **yalğa-* would be expected rather than *alğa-*. As these forms are not confirmed by other languages, it is also unlikely that they represent early 'prebreaking'. It seems most likely that they are recent cases of lowering under the influence of the following vowel, parallel to the developments in EYu *ere-* < **ire-* 'to come', *harβa-* < **hurba-* 'to turn'.

xarğa- ~ xərğa-	**kïrga-*	to shave
alğa- ~ əlğa-	**ïlga-*	to choose
xalğasən	**kïlgasun*	coarse hair

3.6.5. Shirongol evidence for palatal breaking

In a limited set of words the Shirongol languages agree that breaking took place, so that it can be posited for Proto Shirongol. In other lexemes, the Shirongol languages disagree. As in Eastern Yugur, there are also many forms that are inconclusive because of secondary developments of the first syllable vowel.

Words with disagreement among the languages include MgrH *maxa* and Dgx *miğa* 'meat', which makes it hard to judge shapes like MgrM *muqa*, which technically could stem either from the original **mïkan* or from broken **makan*.

In several cases of disagreement between Mongghul and Mangghuer, the latter seems to have the historically correct one, suggesting that Mongghul has a secondary development > *i*, as in MgrH *ćisə*, MgrM *ɕuʒï* < Proto Shirongol **čusun* < CM **čïsun* 'blood', MgrH *ʑirge*, MgrM *ʐurgi* < Proto Shirongol **ǰürgen* 'heart'.[154] As these Mongghul words developed *i* from other vowels, other forms with *-i-* following a palatal consonant, such as *ʑida:* < *ǰïda* 'spear', *śira* < *sïra* 'yellow', *ćire:* < *čïraï* 'face', can not be used as evidence *against* breaking in Mongghul. Whether these forms with *-i-* go back to forms with breaking, can only be established if the corresponding Mangghuer form preserves the broken vowel.

Rounding of **ï/*i* can not with certainty be attributed to breaking when the word starts with a labial consonant, as in MgrH *buru*: < **bïrau* 'calf', *budən* < **bidüün* 'coarse', Dgx *bunʑu ~ bənʑu* < **bïljïur* 'bird', as these could be 'normal' labialisations that may have taken place in any period. The same uncertainty occurs in some words with a labial element following the *i*, as in MgrM *ʂubigi* < **sibüge* 'awl', where the *-u-* may be due to the *b* rather than to breaking. MgrM *ʂudu* 'tooth' probably is a secondary development of *ʂdu* < **sidün*, motivated by the reduced tolerance for initial clusters.

[154] Cf. also MgrH *śilo:*, MgrM *ʂuli* < **silön* 'soup', which has a complicated phonetic history.

Absence of breaking in the Monguor languages can be demonstrated in words originally starting with **kï/*ki*, because these developed initial *ć-*. This applies to MgrH *ćirval* < **kïlbar* 'easy', as well as the reflexes of **kïmusun* 'nail, hoof', **kïrga-* 'to shave', **kïtad* 'Chinese', **kiröe* 'saw', etc. In this regard the Monguor languages stand apart from both Eastern Yugur and remaining Shirongol (the Baoanic languages). Only Dgx *ćirəu* 'saw' inexplicably also features initial *ć*, while normally **kï/*ki* results in *qə*/*kə* there.[155]

Among the words with initial consonants that are neither labial nor palatal, there are a number which feature breaking in several Shirongol languages.

MgrH	Dgx	Kgj	CM	
maxa	*miğa*	*mağa*	**mïkan*	meat
nuġ-la-	*nuqu-*	*nᵾġe-*	**nïku-*	to knead
nəmpəsə	*nuğusuŋ*	*nᵾrsᵾn*	**nïlbusun*	tear
nudurġa	---	*nudurğa*	**nïdurga*	fist
nurə	*nuruŋ*	*nuru*	**nïruun*	back
nədə-	*nudu-*	*nᵾdᵾ-*	**nidü-*	to pound
nudu	*nudu*	*nᵾdᵾ*	**nidün*	eye

Cases supported by both Monguor languages include:

Mongghul	Mangghuer	CM	
ćino ~ ćuna:	*ɕuna*	**čïno*	wolf
ʒ́ulo:n ~ ʒ́ilo:n	---	**gïloan/*jïloan*	smooth
ʒ́ülʒə	*ʑursï*	**jïlsun*	glue
ʒ́oro	---	**jïroa*	ambler
ʒ́u:rə-	*ʑuri-*	**jïru-*	to paint
narġai	*nʲarġai*	**nïlka X *nïraï*	young and tender
uldə	*ulʲdu*	**ildü*	sword
śolgo	*ʂurke*	**silükei*	saliva
śüro:- ~ śiro:-	---	**hiröe-*	to bless
śirən ~ śürun	*ʂuruŋ*	**sirüün*	coarse
ćudor	---	**čidör*	hobble

MgrH *ćuna:,* MgrM *ɕuna* < **čïno* 'wolf' is peculiar in that the second syllable lost its rounding after assimilating the first vowel (cf. EYu *čəna*, Bao *ćina*, where the rounded vowel left no trace at all).

[155] This seems to suggest borrowing from a Monguor-type language, although such a relation between Dongxiang and Monguoric is not otherwise known to exist.

In other cases the Monguor languages disagree with each other, as in MgrH *ćisə*, MgrM *ɕuʒï* < **čïsun* 'blood'. As mentioned above, it is most likely that the Mongghul form is due to 'repalatalisation' of an earlier Monguoric form **ćusə* due to the initial consonant.

Within Baoanic there is little agreement concerning palatal breaking, but the following cases are supported by at least one language:

BaoÑ	Kangjia	Dongxiang	CM	
ćaŋle-	*čiauli-/čiɔli-*	*ɕanlie-*	**čïŋla-*	to listen[156]
---	---	*sanʑa*	**sïlǰa*	sheep tick
ćisəŋ	*čisun*	*ɕusuŋ*	**čïsun*	blood
---	---	*ʂuluŋ*	**sïluun*	upright
ćələ	*čələ: ~ čilə:*	---	**čilöe*	free time
kuru	*kirʉ*	*ćirəu*	**kiröe*	saw
niʂgaŋ	*nišiliʉ*	*muɕuğuŋ*	**ničügün*	naked
nədə-	*nʉdʉ-*	*nudu-*	**nidü-*	to pound
nədəŋ	*nʉdʉ*	*nuduŋ*	**nidün*	eye

When looking at the QG forms of **kïrga-* 'to shave', alternations like EYu *χarğa- ~ χərğa- ~ χurğa-* and Kgj *qarğa- ~ qerğa-* again confirm the instability (and limited usefulness for historical purposes) of first syllable vowels. Given the distribution of forms that look broken and those that do not, and the alternations within languages or groups (e.g. BaoÑ *χarġa-*, Dgx *ġïğa-*), it is most likely that Proto Yugur, Proto Monguor and Proto Baoanic all had preserved high vowels. In some cases there is evidence for recent breaking. For instance MgrM *ćarġa-*, as opposed to MgrH *ćirġa-*, shows that the *a* in the first syllable is probably due to a recent assimilation to the second syllable, because an early assimilation would have prevented the palatalisation of initial **k-*. Occasionally there is a clear indication that breaking must have taken place quite early. Dgx *saŋʑa* < **sïlǰa* 'sheep tick' must have developed its first *a* early enough to prevent the palatalisation of the initial *s-*.

3.7. Development of simple **o*

3.7.1. Default development

In monosyllables and the accented syllable of longer stems **o* generally stays intact in Dagur and Eastern Yugur. Within Shirongol only Kangjia retains **o* as *ɔ*. In remaining Shirongol the standard reflex is *o* (with the usual diphthongal realisations), which has merged with its harmonic counterpart **ö*. Examples include **ol-* 'to find',**ora-* 'to enter', **hodun* 'star', **kola* 'far',**morïn* 'horse'.

[156] BaoÑ *ćaŋle-* may also owe its *a* to the influence of the following *-ŋ-*, see 3.4.4.

Labial breaking of the sequence **o-a* in Dagur was touched upon in 3.2.2. above. It occurs frequently, but not systematically, cf. *war-* < **ora-* 'to enter', but *xɔl* < **kola* 'far'. It is not understood why, although the usual suspicion of central Mongolic and/or Manchu influence applies here as well.

3.7.2. Raising

CM words with the sequence **o-a*, developed a *ɔ-ɔ* (Shirongol *o-o*) sequence due to harmonic rounding of **a* (see above). The accented secondary *ɔ* of the non-first syllables is generally better preserved than the orginal *ɔ* of the first syllable.

In Mongghul, words of this structure tend to neutralise the first (unaccented) *ɔ* to *ə* or raise it to *u*, resulting in forms such as *ulon* < *olon* < **olan* 'many', *sunosə-* < *sonosə-* < **sonas-* 'to hear', *ləmoːn* < **lomoːn* < **noman* 'mole (animal)'.

As mentioned above in 3.2.2., Kangjia tends to develop a vowel sequence *u-u* in words of this structure. Although the Mongghul and Kangjia developments came about independently after harmonic rounding took place, they are both illustrated in the following table, compared with the more primitive situation in Eastern Yugur:

Mongghul	Kangjia	E. Yugur	CM	
ulon	*ulu*	*ɔlɔn*	**olan*	many
uro-	*uru-*	*ɔrɔ-*	**ora-*	to enter
χulo	*ġulu*	*χɔlɔ*	**kola*	far

The development of unaccented **o* > *u* is not uncommon in Kangjia, even in words that did not develop the *u-u* vowel sequence, such as *suqta-* < **sogta-* 'to get drunk', *sunsï-* < **sonas-* 'to hear'. The development **o* > *ʉ* is also found, e.g. *mʉrtʉn* < **modun* 'wood', *fʉtɔ* < **hodun* 'star', although Kangjia *ʉ* more typically represents CM **ü*.

The development **o* > *u* (or *ʊ*) in Mongghul and Mangghuer can also be triggered by adjacent uvular consonants, as in **konï* 'sheep', **mogaï* 'snake', **nokaï* 'dog'.

In Eastern Yugur the raising of **o* > *u* is mainly seen in long vowels, but it also happened in the secondary vowel length of *muːdən* (*mɔːdən* in Qinglong dialect) < **modun* 'wood'. This does not seem to be an Ordos-type raising under the influence of the following high vowel.[157]

3.7.3. Unrounding

Many cases of unrounding of **o* cannot be separated from other phenomena such as raising and devoicing. However, in Shirongol there are instances of unrounding in unreduced, and even in accented syllables. The result is often *a*, as in MgrH (rare)

[157] As suggested by Rybatzki 2003:370.

basə- < **bos-* ‘to get up’, *mańɔǥ* < **mončag* ‘crest’, MgrM *lan* < **nom* ‘book’, Kgj *bağəni ~ bɔğəni ~ bəğəni* < **boganï* ‘low’, BaoD *tabćiə* < **tobčï* ‘button’. The Mongghul reflexes in *fœn* < **hon* ‘year’, *teśʒ́ə* < **tobčï* ‘button’ can be viewed as further developments of *a*.

In most of these cases, the preceding or following labial consonant may have played a role. Unrounding is frequent in Dongshan Mongghul, often after a labial initial consonant, as in *mazə* < **morïn* ‘horse’, *bara* < **bora* ‘grey’, *pasə-* < **bos-* ‘to rise’, *bal-* < **bol-* ‘to ripen’, but also after other initials, as in *naşdo:* < **nogta* ‘halter’, *nayo:n* < **noyan* ‘lord’, *salġʊi* < **solagaï* ‘left’, *talġʊi* < **tolagaï* ‘head’.

In Dahejia Baoan word-final *ɔ* (stemming from **a* by harmonic rounding) can result in *ə* (as can the other non-high vowels), as in *orə-* < **oro-* < **ora-* ‘to enter’.

In EYu *čəna* < **čïno* ‘wolf’ the second syllable was unrounded before palatal breaking could take place.

3.7.4. Palatalisation

Palatalisation of **o* is generally caused by preceding palatal consonants; **y* can also palatalise a preceding **o*. In the QG languages it is hard to separate it from raising and reduction. This is especially true for the Shirongol languages, which lack a phoneme *ö*. Therefore the result of palatalisation of **o* in these languages is mostly *i* or *ü* (the latter is present in most dialects as an allophone or a loan phoneme).

Eastern Yugur has various reflexes, including *ö*, as in EYu *sɔyɔ: ~ söyɔ: ~ sio* < **soyaa* ‘tusk; stalk’, *niyɔn ~ nyɔ:n ~ nion* < **noyan* ‘lord’.[158]

Mongghul examples include *ćugu- ~ ćigu-* < **čokï-* ‘to peck’, and due to initial consonant strengthening also *ćugu- ~ ćigu-* < **ǰokï-* ‘to suit’.

An example with retained rounding from Kangjia is *ǰʉčɔ ~ ǰüčɔ* < **ǰočïn* ‘guest’, and one with collapse of the first syllable *niɔ* < **noyan* ‘lord’.

3.7.5. Elision

Elision of first-syllable **o* will not affect the harmonic rounding of any following **a*, cf. EYu *hġɔr*, BaoD *ġor*, Kgj *χɔr ~ χuar* < **hokar* ‘short’.

Elision of **o* via reduction

E. Yugur	*ru:n*	**oran*	place
Mongghul	*losə*	**olasun*	hemp
	yo:-	**oya-*	to sew
	ʒ́oġʊi	**oǰagaï*	penis
Ñantoq Baoan	*ləŋ*	**olan*	many

[158] Cf. also the palatalisation of diphthong elements, as in EYu *öiro* < **oïra* ‘nearby’.

Elision of **o* via devoicing

E. Yugur	*htɔrɔ*	**dotara*	inside
	qutul-	**hogtal-*	to fell
Mongghul	*ʂdolə-*	**hogtal-*	to fell
	χǧuar	**hokar*	short
Dahejia Baoan	*ġor*	**hokar*	short
	tχuŋ	**togaan*	pot

3.8. Development of simple **ö*

3.8.1. Default development

In Dagur, CM **ö* has normally merged with **ü* into *u*, as in *bur*^{y} < **bös* 'fabric', *nuɣur* < **nöker* 'friend', *kuk*^{w} < **köke* 'blue'.

In Eastern Yugur, CM **ö* has resulted in a real front vowel *ö*, which in many words has become centralised to closed *o* (which can be distinguished from *ɔ*, at least in Bolčuluu's analysis). However, closed *o* could also represent an earlier stage of development than the front *ö* (cf. the development of **ü* below).

In Kangjia the distinction between **ö* and **o* was preserved. In the remaining Shirongol languages **ö* merged with its harmonic counterpart **o* into the modern phoneme *o*. This default development is generally found in monosyllables such as **ög-* 'to give', **bös* 'fabric', **köl* 'foot'.

The same reflexes may be found in longer stems. However, originally **ö* was mostly restricted to the first syllable, apart from instances of harmonic rounding of **e*, e.g. MgrH *losə-* < **olos-* < CM **öles-* 'to be hungry', and in words with **ö* in the second syllable following *i* of the first syllable, e.g. **čidör* 'hobble', **silön* 'soup'. The latter category is the only one with second syllable **ö* that can not be explained as secondary development. The existence of **ö* after **e* or **ü* of the first syllable cannot be demonstrated for older Mongolic based on the available evidence.

Two common deviations from these developments are unrounding and palatalisation.

3.8.2. Unrounding

Unrounding of **ö* in first syllable is unpredictable but not uncommon in QG. In Eastern Yugur most examples of unrounding seem to be dissimilation cases due to a preceding labial consonant, e.g. *mere:n* < **mören* 'river', *mer* < **mör* 'path', *pes* < **bös* 'fabric'. A variant with *ö* is also attested in all these cases. It may be relevant that these words contain labial consonants. However, the *ö* variant of *A*-suffixes (*A* stands for the fourfold alternation *a e ɔ ö*) tends to be replaced by the *e* variant, irrespective of the surrounding consonants.[159]

[159] In Western Yugur there is a similar tendency to unround **ö* in the vicinity of a labial consonant, as in *perïk* < **bö:rk* 'hat'.

All Shirongol languages have some examples for unrounding of **ö*, e.g. MgrH *far ~ for ~ fod* < **höd(ü)* 'larva', Dgx *kəwaŋ* < **köbeün* 'son', Kgj *besge* < **bögse* 'buttocks', *mante-* < **möltül-* 'to undo'. Although many of the affected lexemes contain labial consonants and thus may involve dissimilation, other cases are unpredictable, e.g. MgrM *aŋgo* < **öŋge* 'colour', Dgx *śieni*, Kgj *sïni* < **söni* 'night', Dgx *śinie-*, Kgj *sïni-* < **söne-* 'to extinguish', Kgj *tere-* < **töre-* 'to give birth', Kgj *eteɻ-* < **ötel-* 'to grow old'.

A few words feature unrounding in several languages. For some words it may be viewed as an older development in view of their distribution. **dörben* 'four' is unrounded in all of Shirongol and may therefore have been unrounded in Proto Shirongol (cf. also MgrH *təʑin* < **döčïn* 'forty', which was lost in remaining Shirongol). The same applies to **möŋgün* 'silver'.

In Baoan and Dongxiang we find unrounding in **mölsün* 'ice' and **sölsün* 'gall bladder'. The fact that this unrounding did not occur in Kangjia means that it is not a development of the Proto Baoanic period.

In Dahejia Baoan word-final *o* < **ö* (< **e* by harmonic rounding) can result in *ə*, as in *orə* < **öre* '(pit of the) stomach'.

Second syllable **ö* following **i* of the first syllable was unrounded in EYu *čüder* < **čidör* 'hobble'. EYu *šelen ~ šölön* < **silön* 'soup'. See the section on palatal breaking in 3.6. above.

The unrounding in Dag *bədən*^{y} < **bödene* 'quail' and *bəlluɣ*^{w} < **böldegen* 'testicles' is due to the labial initial.

3.8.3. Raising

In Eastern Yugur the reflex *ᵻ* is not uncommon in Hongshiwo dialect (the basis of Bolčuluu's dictionary), as in *mᵻren* < **mören* 'river', *hᵻtgwen* < **ödken*, but in Qinglong dialect **ö* is apparently retained. However, in view of forms such as Qinglong *nödön* < **nidün* 'eye', at least some occurrences of *ö* may in fact be due to lowering of an earlier development **ü*.

In Mongghul the sequence **ö-e* was harmonically rounded > *o-o*, and has the tendency to develop into *u-o* (like its harmonic counterpart **o-a* > *o-o*), e.g. **töre-* > *turo-* 'to be born'.

In parallel to **o-a*, Kangjia has the development of the sequence **ö-e* > *ö-ö*, which may be raised to a sequence *ᵻ-ᵻ*, apparently mostly in the vicinity of velars. It is not clear whether this *ᵻ-ᵻ* developed via an intermediate stage whereby only the first or the second **ö* was raised.

Some Mongghul and Kangjia cases compared:

Mongghul	Kangjia	E. Yugur	CM	
kugo-	*kᵻkᵻ-*	*hkö-*	**köke*	blue
nuko	*nᵻχgᵻ(ŋ)*	*nökön*	**nöken*	cave
sgo:-	*sᵻgᵻ-*	---	**sögee-*	to scold

Kangjia *ʁ* can also be the result of **ö* in other environments, *e.g. ʁɣ-* < **ög-* ‘to give’, *ʁndər* < **(h)öndür* ‘high’, *ʁle* < **öreele* ‘half of a pair’, *ʁramɔ* < **öre aman* ‘pit of the stomach’.

3.8.4. Elision

Elision of **ö* via reduction

E. Yugur	*lögčən*	**ölegčin*	she-dog
	nöčən	**önečin*	orphan
Mongghul	*losə-*	**öles-*	to be hungry
	ŋgo	**öŋge*	colour
	ndur	**öndür*	high
	rme:n	**öreme*	skin on milk
Kangjia	*ŋgʁsʁn ~ gʁsʁn*	**örgesün*	thorn
	lesï-	**öles-*	to be hungry
Dongxiang	*gieru-*	**öŋgere-*	to pass by
	muğə	**örmege*	coarse fabric

Elision of **ö* via devoicing

E. Yugur	*hkö*	**köken*	blue
	hkö-	**köke-*	to suckle
	sʁkel- ~ sgöl-	**öskel-*	to kick
Mongghul	*ʂdogo:n*	**ödken*	dense
Mangghuer	*diməi*	**ödme*	bread
	xu-	**ög-*	to give
Ñantoq Baoan	*ʂgel-*	**öskel-*	to kick
Kangjia	*dagɔ ~ ʒïgɔ*	**ödken*	dense

**öčügedür* ‘yesterday’ has lost its first vowel in Eastern Yugur and in all of Shirongol with the exception of Dongxiang: EYu *čogdor*, MgrH *ćigudur*, MgrM *ʈugodur*, BaoD *gudər*, Kgj *šigʁdə*, Dgx *fuʒuğudu*. Remarkably, initial vowel loss is even found in this word outside of the QG languages, as in Ordos *čʁgdʁr* and Moghol *čikaʊdur*.

Another example of initial vowel loss that affects Mongolic languages both in the QG area and elsewhere is **öčeen ~ *čöen* ‘few’, if the resemblance between the two forms is not a coincidence. The loss of the initial vowel of **ösegei* ‘heel’ is restricted to Buriat.

3.9. Development of simple **u*

3.9.1. Default development

In Dagur **u* generally merged with **o*, typically resulting in *ɔ*. In Shirongol (except Kangjia) it merged with **ü*. In Kangjia it was usually retained, and can be distinguished from its harmonic counterpart **ü*, although fluctuations occur. In neutral consonant environments, the default reflex of **u* is *u*, e.g. in **gurban* 'three', **kurïm* 'banquet', **sula* 'loose', **sun-* 'to extend', **unagan* 'foal'.

In Eastern Yugur **u* is also generally retained as *u* (in Bolčuluu's analysis). There are some alternations between *u* and *ɔ*, eg *yɔğɔr ~ yuğur* 'Yugur', *ɔrɔġ ~ urəġ* < **urug* 'relatives by marriage'. Occasionally medial **u* was lowered in Shirongol by **a* of the following syllable, most notably in MgrM *ćidoġo*, Dgx *qudoğo*, from **kïtogo* < **kïtoga* < **kïtuga* 'knife'. Cf. also Dgx *yidoğo* < **ïdugan* 'shaman'. This is reminiscent of the development of the complex vowel **oa* < **ua* in **ǰïloa* 'reins' and the like.

3.9.2. Unrounding phenomena

High vowel neutralisation

The treatment of **u* in accented syllables varies. It is often preserved in Baoanic, but usually unrounded to *ə* in Monguoric and Eastern Yugur, as can be seen in words with the suffixes *-sUn*, *-dUn*, *-tU*, in Eastern Yugur also (verbalizer) *-dU*, (nomen futuri) *-kU*, and (deverbal noun) *-bUrI*. The *U* of course represents **ü* as well, which in these suffixes results in the same sounds in the QG languages (Kangjia excepted). This can be illustrated by means of **yasun* 'bone': EYu *yasən*, MgrH *yasə*, MgrM *yaġsï*, BaoÑ *yasəŋ*, Kgj *yasun*, Dgx *yasuŋ*.[160]

In Eastern Yugur and Mongghul the high vowels **u* and **ü* as well as unrounded **ï* and **i* may all result in *ə*, effectively leading to their neutralisation. Although *ə* is phonetically a mid vowel, it remains in the phonological high vowel category, not only for historical reasons, but also because of the alternations between *ə* and *i* and *u* in the modern languages.

In longer stems with more than a single *u*, they apparently strengthen each other's presence, and resist the neutralisation > *ə*, cf. EYu *qusun* < **usun* 'water', *nurğusun* < **nugursun* 'spinal marrow', *χulusun ~ χulusən* < **kulusun* 'reed', *urus-* < **urus-* 'to flow'. In some words only one of the original **u*'s was preserved, but nevertheless the expected unrounding did not take place, e.g. MgrH *sʒu* < **usun* 'water', EYu *hdu-* < **hutu-* 'to emit smoke'. The **u*'s of stems with other vowels in the first syllable, such as **nasun* 'age', **daru-* 'to press', **modun* 'wood', were unrounded in Eastern Yugur and Mongghul.

In the initial syllable the development **u* > *ə* rarely occurs in non-labial environments, as in EYu *ǰəlǰəğan* < **ǰulǰaga(n)* 'animal young'.

[160] **nasun* 'year of age' resulted in Baoan *nasə*, apparently because final *-n* was lost.

In the vicinity of labial consonants it is impossible to specifically attribute examples of unrounding of **u* to dissimilation or to high vowel neutralisation. Examples from Eastern Yugur include *bəda:n* < **budaan* 'porridge', *pərćaġ* < **burčag* 'bean', *səmən* < **sumun* 'arrow', *nəmən* < **numun* 'bow', and *səb* < **sub* 'otter'. There are also examples from Shirongol, e.g. BaoÑ *bəla-* < **bula-* 'to bury', Kgj *bəlaŋgar ~ bʉlaŋgar* < **bulaŋgïr* 'turbid, unclear', Dgx *bili-* < **bulï-* 'to snatch', Dgx *pïʒa ~ puʒa* < **burčag* 'bean'.

There must have been prehistoric alternations **u ~*ï* near labial consonants, partly documented in Middle Mongol, see **bulčïrkaï* 'gland', **bïrau* 'calf', **busï* 'not'. Given that the **h-* in**huǰaur ~ *hïǰaur* 'root' probably goes back to PM **p-*, the vowel alternation could also be ascribed to this labial consonant.

Two rare Dagur examples for this type of unrounding are *bəlʲʸe:-* < **bulïa-* 'to snatch', and *xənʸe:s* < **kunïasun* 'fold' (the expected forms with **ɔ* also exist).

Labial breaking

Labial breaking in Dagur involves the insertion of second syllable **a* into a first syllable with **u* or **o*, while this original labial is reduced in the process, and the **a* usually disappears from its original position, e.g. *dʷand* < *dŭandă* < **dumda* 'middle'. Further examples include *want-* < **umta-* 'to sleep', *gʷay* < **guya* 'thigh', *xʷark* < **huraka* 'lasso', *sʷadəl* < **sudal* 'vein'. Also in this regard **u* behaves just like **ɔ* in Dagur, cf. *war-* < **ora-* 'to enter', *dʷatər* < **dotar* 'inside'.

After a labial consonant the vowel loses the labial element, or put differently, the labial initial consonant loses its labial coarticulation: Dag *baɣ* < **buka* 'bull', *bačil-* < **bučal-* 'to boil', *manə* < **muna* 'cudgel'. It is not clear whether there ever were intermediate forms with *bʷ-*, e.g. **bʷaɣ* as an earlier shape of *baɣ*.

Despite being an innovative Dagur feature, labial breaking contributes to CM reconstruction, by showing that words with labial breaking must have had **a* in their second syllable. For instance Dag *nʷaɣəs* 'duck' suggests CM **nogasun* or **nugasun*, contrary to Ordos *nuġusu*, which favours **u* in the middle syllable.

Unrounding and lowering

The QG languages also feature developments that are similar to, but independent from, that seen in Dag *baɣ* < **buka* 'bull'. This development **u* > *a* appears to be most frequent after a labial initial, but can occasionally be found elsewhere, as in EYu *balağan* < **bulagan* 'sable', MgrM *barćiġai* < **bulčïrkaï* 'gland', BaoD *balaġ ~ bulaġ* < **bulag* 'source', Dgx *bağaçəu* < **buka(-čug)* 'bull'. However, incidental cases such as MgrH *xɔnaġla ~ xanaġla* < **kulugana* 'mouse', and variants such as EYu *χuyaġ ~ χayaġ* < **kuyag* 'armour', *harβa-* < **hurba-* 'to turn', *haya-* < **huya-* 'to tie' indicate that assimilation to the vowel of the following syllable may play a role here rather than the labial consonant. In Eastern Yugur this type of lowering also occurs in the other high vowels of the first syllable (see for instance 3.5.3.).

Lowering accompanied by inexplicable fronting is seen in BaoÑ *terma* < **turma* 'turnip', BaoÑ *deraŋ* < **duran* 'liking', Kgj *mešχe-* < **muškï-* 'to twist'.

3.9.3. Palatalisation

In the QG languages palatalisation of **u* often co-occurs with unrounding. In the Shirongol languages this is partly due to the lack of a phoneme *ü*. Complete palatalisation **u* > *ə* > *i* by preceding palatal consonants is seen in Eastern Yugur and Shirongol. A widespread example is **ayu-* 'to fear', which has mostly become *ayi-* or *ai-*. A following **y-* can also palatalise **u*, as in EYu *ɣ́a:-* ~ *hia-* < **huya-* 'to tie'. Palatalisation by preceding palatals is nearly universally applied in Mongghul and Baoan, while it is rare in their respective sister languages Mangghuer and Dongxiang, as illustrated by MgrH *ʒ́ila:*, MgrM *ʑula*, BaoÑ *ʒ́ila*, Dgx *ʑula* < **ǰula* 'lamp'.[161]

As *ə* and *i* tend to be in an allophonic relationship in most of Shirongol, developments **u* > *i* in the absence of palatal consonants such as MgrM *bili-* ~ *buli-* < **bulï-* 'to seize', can probably be interpreted as dissimilations.

Palatalisation > *ü* by adjacent consonant is found in Eastern Yugur, e.g. *čüsaġ* (~ *čisaġ*) < **ǰusag* 'two-year-old (sheep)'. The Eastern Yugur phonological system already contained an *ü*, unlike that of Shirongol. Occasional cases occur elsewhere, e.g. BaoD *ġanʒ́üŋ*, Kgj *ġanǰün* ~ *ġanǰun* < **kamčun* 'sleeve'.

3.9.4. Elision

Elision of **u* via reduction

E. Yugur	*na:-*	**una-*	to descend
	βəǰə	**ubuǰï*	feeding bottle
	χwa:-	**kubïa-*	to divide
	sβəd	**subud*	pearl
Mongghul	*ŋġua:-*	**ugaa-*	to wash
	lus	**ulus*	nation
	da:ldə-	**kudaldu-*	to sell
Ñantoq Baoan	*ta:-*	**unta-*	to sleep

A widespread example of initial vowel loss is **uŋsï-* ~ **umsï-* 'to read', which lost its initial vowel in Monguoric and in most Baoan dialects, e.g. MgrH *mośə-*. BaoD *muśi-*, BaoGt *məʂ-*. Elision of the **u* may have become possible after a form **umuši-* with epenthetic vowel had developed. However, due to 'dissenting' forms like BaoÑ *ɵmći-* and Kgj *unči-*, it can not be reconstructed for Proto Shirongol.

**marta-* (or **umarta-*, based on its spelling) 'to forget' which famously lost its initial vowel even in central Mongolic, does not preserve a trace of it in the QG languages either.

---

[161] Secondary *u* that was the result of palatal breaking, e.g. Proto Shirongol **čusun* < **čïsun* 'blood', has undergone the same palatalisation in MgrH *ćisə*, BaoD *ćisuŋ*, as opposed to the more conservative MgrM *ɕuʒi*, Dgx *ɕusuŋ*.

The exceptional loss of initial **u* in Dag *wa:-* < **ugaa-* ‘to wash’ probably involves an intermediate form **uwa:-*.

Elision of **u* via devoicing

E. Yugur	*hta:*	**hutaa*	smoke
	hta-	**duta-*	to lack
	rdə, hurtu	**urtu*	long
Mongghul	*ʂdur*	**urtu*	long
	śʒüra:-	**učïra-*	to meet
Ñantoq Baoan	*hʑa-*	**kuča-*	to bark
	ʂdasəŋ	**hutasun*	thread
Dahejia Baoan	*ġul-*	?**kukul-*	to break (tr.)[162]

**usun* ‘water’ lost its initial vowel in most of Shirongol, e.g. MgrH *ʂʒu ~ sʒu,* BaoD *sə*, Kgj *sʉ ~ sü*. Older sources suggest that EYu *qusun* also goes back to a form with elided vowel *χsun* (*χ-* being a relic of the devoiced first syllable). MgrM *suʒu* also involves a modern development from an older monosyllabic form *sʒu*, due to decreasing tolerance for initial clusters.

Loss of **u* in other syllables can be seen occasionally, in Eastern Yugur usually following *-r-*, cf. *a:r* < **aru* ‘back’, *dar- ~ darə-* < **daru-* ‘to press’, *jar-* < **ǰaru-* ‘to use’.[163] Cases of *Mittelsilbenschwund* include EYu *χunlaġ* < **kulugana* ‘mouse’, BaoÑ *ġəlsəŋ* < **kulusun* ‘reed’.

3.10. Development of simple **ü*

3.10.1. Default development

In Dagur **ü* merged with **ö* into *u* (See 3.8.).

In Eastern Yugur CM **ü* is represented by central *ʉ* (according to Bolčuluu), and in some words by the front vowel *ü*.

In Kangjia the distinction between **ü* and **u* was preserved, although many irregularities occur. In the remaining Shirongol languages **ü* merged with its harmonic counterpart **u* into the modern phoneme *u*. This default development is found in words such as **bürge* ‘flea’, **büse* ‘belt’, **dülei* ‘deaf’, **ǰüg* ‘direction’, **kür-* ‘to reach’, **üge* ‘word’.

As far as the Shirongol languages are concerned, **ü* can also be reconstructed in **übül* ‘winter’ and **üdür* ‘day’, whose Written Mongol spellings suggest original forms **ebül* and **edür*. In these cases the forms with **ü-* are also

---

[162] BaoD *ġul-* did not lose its first syllable by devoicing, if it goes back to the variant **kugul-*.

[163] A rare eample after another consonant is EYu *bat* < **batu* ‘strong’. The loss of the final vowel of the suffixes *-kU* and *-ǰU* is not universal, but only occurs when certain auxiliaries or negative particles follow.

supported by Dagur. In other words with this vowel pattern Dagur is in agreement with the Written Mongol spelling (cf. **elgü-* 'to hang', **ergü-* 'to lift', etc). In **sibüge* 'awl' and **čimügen* 'marrow' it seems better to reconstruct **ö* rather than **ü*; **ö* may be a secondary development of **e* due to the labial consonant.[164]

3.10.2. Unrounding phenomena

High vowel neutralisation

Like its harmonic counterpart **u*, **ü* tends to become *ə* in the final syllable in Eastern Yugur and Mongghul, cf. EYu *mendə* < **mendü* 'peace', EYu *xisən*, MgrH *ɣɜən* < **ye(r)sün* 'nine', EYu *hodən*, MgrH *fo:də* < **hödün* 'feather'. There is also the exception, as in the development of **u*, that words with several **ü*'s may withstand the delabialisation tendency, as in EYu *nʉdʉn*, MgrH *nudu*, BaoÑ *nədəŋ*, Kgj *nʉdʉ (*< **nʉdʉn)* < **nidün* 'eye'. Even when the first syllable vowel is elided, the rounding remains in the second vowel, as in EYu *hsʉn*, MgrH *ɣɜu* < **hüsün* 'hair', EYu *hsʉn* < **üsün* 'milk', EYu *hkʉ-* < **ükü-* 'to die'.

Some of the abovementioned words with the vowel sequence **e-ü* have lowered vowels in Eastern Yugur, reminiscent of central Mongolic, whereas Shirongol has two high vowels. However, this is another issue where the descriptions of Bolčuluu and Junast differ. Eastern Yugur examples include *odor* (B), *udur* (J) < **üdür* 'day', *olɣo-* (B) < **elgü-* 'to hang', *orɣo-* (B), *urɣə-* (J) < **ergü-* 'to lift', *öröɣö* (B) < **erüke* 'smoke-hole'.

In the vicinity of labial consonants it is impossible to specifically attribute examples of unrounding of **ü* to dissimilation or to high vowel neutralisation.

Moreover, there are several words with prehistoric alternations **ü* ~ **i*, as recorded in Middle Mongol and modern languages including **bütegei (?*bitegei)* 'do not', **büteü (?*biteü)* 'double handful', **hüleü* ~ **hileü* 'superfluous', **ildü* ~ **üldü* 'sword'. It is not always clear which variant is the older. The reason behind these alternations is not the same in all cases. In the words starting with **b-*, this labial consonant may have been the trigger to labialise or delabialise the first vowel. As **h-* stems from PM **p-*, one of the variants **hüleü* ~ **hileü* may also be due to assimilation or dissimilation. **üldü* 'sword' can be viewed as an early 'prebroken' variant of **ildü*.

Modern cases of **ü* > *ə* (predominantly near labial consonants) include EYu *bələmbər* < **bülien.bir* 'lukewarm', *pʉsəɣʉi* ~ *pəsəɣui* < **büsegüi* 'woman', BaoÑ *təlgə-* < **tülki-* 'to push'.

Unrounding and lowering

Unrounding and lowering in Eastern Yugur occurs in stems with **ü* in the first and **e* in the second syllable, e.g. *eǰe-* < **üǰe-* 'to see', *hele-* < **hüle-* 'to remain', *tele-* < **tüle-* 'to burn', *temen* < **tümen* 'ten thousand'. Similar developments can be

[164] Khalkha, Ordos and Kalmuck suggest **siböge*, in which the **ö* could either be the **ü* lowered by the following **e* (cf. **kituga* 'knife' for a parallel back-vocalic case), or an independent labialisation of the middle **e* of **sibege*).

seen in other words with a high vowel followed by a non-high vowel, so these cases may merely involve the assimilation of an unaccented vowels to the accented one in the following syllable.

Similar developments can be found in the Shirongol languages, but there most cases seem to involve dissimilation in the vicinity of a labial consonant. Examples include MgrH *tebśin* < **tübsin* 'level', MgrM *bergesï* < **bürküesün* 'lid', BaoÑ *belġaŋ* < *?*bülie.ken* 'lukewarm', *hergə-* <**hürgü-* 'to be startled', *telaŋ* < **tüli.en* 'firewood', Dgx *bəŋğə* < **bürge* 'flea'.

Unlike its harmonic counterpart **u*, short **ü* does not systematically undergo labial breaking in Dagur. Dag *bəs* < **büse* 'belt' may have developed via **bŭəsĕ*.[165]

3.10.3. Palatalisation

Palatalisation of **ü* by a preceding palatal is occasionally seen in Shirongol, e.g. MgrH *ʒirge*, BaoD *ʒirgə*, Kgj *ǰirge ~ ǰirgi* < **ǰürüken* 'heart', MgrM *kuʒi* < **küčün* 'strength', MgrH *nəćigo:r* < **ničügü.er* 'naked', and (strictly from long **üü*: BaoÑ *uʒir* < **(h)üǰüür* 'tip, end'.

3.10.4. Elision

Elision of **ü* via reduction

E. Yugur	*ni:n*	**ünien*	cow
	βəl	**übül*	winter
	noɣon	**ünügün*	kid (goat)
Mongghul	*sme:n*	**süme*	temple
	dur ~ udur	**üdür*	day
	ʒu:r ~ uʒu:r	**üǰüür*	tip

A widespread case is the negation particle **üle ~ *ülü*, cf. EYu *lə*, MgrH *li:*, BaoD *lə*, Kgj *ne*, and even Mog *lʉ*, *lɛ*, *la*. As this word does not carry an accent, it frequently becomes prefixed to the verb.

Elision of **ü* via devoicing

E. Yugur	*hsʉn*	**hüsün*	hair
	kʉte-	**ügtee-*	to weed
Mongghul	*sʒə ~ fuʒə*	**hüsün*	hair
Ñantoq Baoan	*ʂge*	**süke*	axe
	se	**büse*	belt

[165] Labial breaking in front-vocalic words seems to be essentially limited to **öe*, cf. Dag *xʷə:s* < **höesün* 'pus', *tʷə:* < **töe* 'handspan', *mʷə:r* < **möer* 'rim of a wheel', etc.

Some words have elided **ü* of the first syllable in several QG languages, including **hüsün* 'hair' listed above, **ükü-* 'to die', and **ütergen* 'threshing floor': MgrM *śdərge*, BaoÑ *terġaŋ*, Kgj *tɔrğɔ*, Dgx *tuğoŋ*.

In a couple of words the presence and absence of initial **ü* has an unusual distribution. In **üsü(n)* ~ **sün* 'milk', the former variant is supported by Eastern Yugur as well as by the central languages Kalmuck and Ordos. It seems unlikely that the longer form **üsün* is a secondary extended form. It is potentially very old, and reminiscent of **umarta-* ~ **marta-*, although it lacks the early documentation of that word. Likewise, the word otherwise reconstructed as **gün* 'deep', has an additional syllable in MgrH *fugon* ~ *gun*, suggesting a reconstruction **ügün* or even **öken*, which is however not confirmed by any other language.

Vowel loss in non-first syllables usually involves *Mittelsilbenschwund*. In case of **jürüken* 'heart' this occurred in all of Shirongol and exceptionally also in Eastern Yugur. Other examples include MgrH *śolgo*, MgrM *ʂurke*, Dgx *ʂəŋki* < **silükei* 'saliva', BaoÑ *ebdəg* < **ebüdüg* 'knee'.

3.11. Issues related to vowel length

3.11.1. Vowel length in Mongolic

In this section I will review some of the possible evidence for original vowel length in Eastern Yugur and the Shirongol languages, with some excursions on Dagur. The main focus will be on the determination of the historical accuracy of modern vowel lengths, and on finding any alternative explanations. The synchronic distinctive value of vowel length in the peripheral languages will nevertheless be discussed.

The feature of vowel length is found in many Mongolic languages. The distinction between short and long vowels is phonologically relevant in the central languages Khalkha, Buriat, and Kalmuck, and in Dagur. In the QG languages distinctive vowel length is found in Eastern Yugur, in Mongghul; and it seems to preserve some relevance in the Ñantoq and Xiazhuang dialects of Baoan.

The rare vowel lengths attested in Mangghuer (only indicated by Čenggel-tei) and in Kangjia, Dahejia Baoan and Dongxiang (in the latter only documented by Todaeva) can only be of limited phonological importance synchronically. They are however of historical importance.

Long vowels have been described for Moghol, but these are restricted to certain vowel qualities (under the influence of the Tajik vowel system) and lack phonological significance.

In Dongxiang, Mangghuer, and Moghol the lengths that result from double vowels (and possibly lengths of other origins) have generally merged with the short vowel set. Incidental surviving vowel lengths in these languages are valuable for historical purposes. All languages that lost phonological vowel length did preserve some of the complex vowels in the form of diphthongs.

Examples demonstrating the phonological relevance of vowel length in the various languages will be given below. After that follows a section on the various origins of vowel length, including a discussion of so-called primary vowel length from the viewpoint of Dagur and QG languages. The modern languages are important for the diachronic discussion, as Written Mongol spellings and Middle

Mongol notations are not always informative, especially where those long vowels are concerned that are not the result of contractions.

3.11.2. Phonological value of vowel length

In order to establish the phonological 'weight' of vowel quantity in Mongolic, I will list some 'minimal pairs' and 'near minimal pairs' for each of the peripheral Mongolic languages that preserve long and short vowels.

The importance of such pairs in daily speech must not be overestimated. The opposed words are often different parts of speech so that they are not interchangeable in a given grammatical context. Furthermore, however respectable the list of minimal pairs of a language, the majority of words, whether short or long vocalic, do not have an otherwise identical counterpart of the other quantity. The minimal pairs given below at least demonstrate that the vowel quantity is not determined by the phonetic environment.

Vowel length is always a relative feature. Long and short vowels do not have a fixed duration. Many irregularities and variations within and between languages are probably due to this fact.

Words which have a counterpart probably preserve their quantity better. But even phonemically relevant vowel lengths may disappear. Long vowels then merge with their short counterparts and eventually the entire system of vowel length distinction can be lost.[166] This is what happened in Mongolic, apparently three or four times independently: in Moghol[167], Mangghuer, Dahejia Baoan, and probably separately again in Dongxiang (which seems to descend from a conservative Baoan dialect not unlike Ñantoq Baoan (without the Tibetan element).

The (rather incidental) vowel lengths observed by Todaeva in Dahejia Baoan, and the vowel lengths recorded by Čenggeltei in Minhe Monguor, will be briefly discussed below.

Historically, vowel lengths have various origins. Double vowels **aa*, **ee*, **ïï*, **ii*, **uu*, **üü* gave rise to modern long vowels, naturally with various qualitative shifts similar to their short counterparts. Contraction of complex vowels **oa*, **öe* (as they are reconstructed here) resulted in additional long vowels *o:* and *ö:*.

As expected, some of the minimal or near-minimal pairs in the peripheral languages are direct continuations of ancient simple-double oppositions such as CM **dere* 'pillow' ≠ **deere* 'above'. However, many of the CM minimal pairs of this type did not survive. Most minimal pairs found in Dagur, Mongghul and Eastern Yugur are new. The long vowels may result from the contraction of diphthongs or my have other secondary origins. As seen in the examples below, long vowels in the peripheral languages may correspond to short vowels elsewhere in Mongolic, and short vowels may be shortened long vowels. In many cases one or both members of a minimal pair may be loanwords.

---

[166] Compare the fate of the Common Turkic vowel length, which only survived directly in three modern languages, out of some thirty modern languages.

[167] The modern Moghol vowel lengths usually developed under Persian influence. See Ligeti 1964.

Different descriptions of the same language may disagree about the quantity of a vowel. Some divergences may be due to small dialectal or ideolectal differences, others perhaps merely to different interpretations by the various authors. Junast hears MgrH *xumba-* 'swim', *nura:-* 'collapse', *ʒi:da* 'spear', while Khasbaatar hears *xumba:-*, *nu:ra:-*, *ʒida:* (CM **humba-*, **nura-*, **ǰïda*). Similarly, Enkhbat hears Dag *dəlin* 'udder', *xɔrdun* 'quick', whereas Namcarai & Khaserdeni hear *dəli:n*, *xɔrdɔ:n* (CM **delen*, **kurdun*). Such discrepancies also occur in words which are expected to have long vowels. Enkhbat's *kunšun* 'smell of burning' appears as *kuršu:n* in Namcarai & Khaserdeni (CM ?**küŋsiün*).[168] In the case of Enkhbat's *xɔllə* as opposed to Namcarai & Khaserdeni's *xɔru:l* 'lip' the situation is the reverse, but for this lexeme the existence of two early variants **huruul* ~ **hurul* is confirmed by the other languages. Many Eastern Yugur words have a long vowel according to Junast and a short vowel according to Bolčuluu & Jalsan:

Junast	Bolčuluu & Jalsan	CM	
mere:n	*muren*	**mören*	river
ule: ~ ölö	*ule*	**üile*	work
kude:l-	*kutel-*	**kötel-*	to lead
urǰüno:n	*ɔrǰinɔn*	**urïǰï hon*	year before last

These are all historically short vowels which have apparently been lengthened in the accented final syllable.[169]

Variation in quantity is also commonly found within the materials of one author, e.g. MgrH *dalə ~ dali:* 'alike', *burge ~ bu:rge* 'flea' in Khasbaatar, Dag *ɔlum ~ ɔlɔ:n* 'bellyband' in Enkhbat. The relative abundance of such variations seems to go against the case for vowel length as an important distinction.

3.11.3. Vowel length distinctions in peripheral Mongolic

Apart from the central languages Khalkha, Ordos, Khamnigan, Buriat, and Kalmuck, long and short vowels are distinguished in Dagur, Eastern Yugur, Baoan of Ñantoq, Baoan of Xiazhuang, and Mongghul. The phonological relevance of vowel quantity can be illustrated by means of minimal pairs.

In the central languages, minimal pairs often correspond directly to CM pairs. E.g. the short vowels of Khalkha correspond to short CM vowels, and the long vowels of Khalkha correspond to double vowels.

In Central Mongolic there are numerous minimal pairs in which the word with the long vowel is the causative of the word with the short vowel, e.g. Khalkha *xatax* 'to dry (intransitive)', *xataax* 'to dry (transitive)', from **kata-* and **kata.a-*,

[168] Note that many of these words are also found in Manchu and Northern Tungusic languages. These languages may be the source of variant forms in Dagur or may otherwise have affected the Dagur forms.

[169] The reverse situation seems to be less common, but can be observed, e.g. in Junast *görösən*, Bolčuluu *gure:sən* 'wild animal' (**göreesün*), Junast *ğua-*, Bolčuluu *uğwa:-* 'to wash' (**ugïa-*), Junast *jü:ra-*, Bolčuluu *jüra-* 'to mix' (**ǰuura-*).

respectively. The concept of phonemic vowel length itself was preserved in several peripheral languages, but, due to various sound shifts, many of the original oppositions do not survive as such. Many of the minimal pairs in the peripheral languages have arisen due to incidental, often unpredictable developments. In other cases vowel qualities have merged so that the distinction between two words comes to rely on the vowel quantity, as in the following example from Ñantoq Baoan:

BaoÑ	CM	
ɵːləŋ	**eülen*	cloud
ɵləŋ	**olan*	many

In the following, a selection of clear oppositions between short and long vowels will be presented for each of the peripheral languages where vowel length is still distinctive. For the other languages a selection of relictual vowel lengths will be given.

3.11.3.1. Vowel length distinctions in Dagur

See the lists of Enkhbat 1988:58, Zhong 1982:11, cf. also the vocabulary in Namcarai & Khaserdeni 1983.

Dag	source	
čaːs	**čaasun*	paper
čas	**časun*	snow
baːs	**baasun*	excrement
bas	**basa*	also
saːr	**saarï*	leather strips
sar	**sara*	moon
ǰaːr	**ǰïar*(ï)	musk
ǰar-	**ǰaru-*	to use
əməːl	**emeel*	saddle
əməl	**emüne*	front
tɔːs	**toasun*	dust
tɔs	**tosun*	fat
duːr-	**düüre-*	to fill up (intr.)
dur-	**düre-*	to sell
xuːl-	**hülie-*	to blow
xul-	**hüle-*	to remain

3.11.3.2. Vowel length distinctions in Eastern Yugur

Most examples are listed in Bolčuluu & Jalsan 1990:51-52. When EYu /*i:*/ is interpreted as the long counterpart of /*ə*/, pairs such as *šəra* 'yellow' ≠ *ši:ra* 'leg' can also be viewed as minimal pairs involving vowel length. See Nugteren 2003.

EYu	source	
ǰa:r	**ǰïar(ï)*	musk
ǰar-	**ǰaru-*	to use
ġada:r	**kadaar*	bridle
ġadar	**gadar*	surface
sa:r	**saarï*	buttocks (of livestock)
sar	**sar*	bird of prey
da:rə	**daarï*	saddle gall
darə-	**daru-*	to press
ta:-	**taa-*	to guess
ta	**ta*	you (pl)
ta:wən	**tabun*	five
tawən	**tabïn*	fifty
xana:-	**kanïa-*	to cough
xana	**kaana*	where
xwa:r	**kabar*	nose
xwar	Amdo *xwar*	printing block
de:re (~ *di:re*)	**deere*	above
dere	**dere*	pillow
de:l	**deel*	garment
del	**del*	mane
teme:n	**temeen*	camel
temen	**tümen*	ten thousand
ɔrɔ:	**oraa*	untamed
ɔrɔ-	**ora-*	to enter
xɔrɔ:-	**horïa-*	to wind
xɔrɔ	**kora*	poison

tölö:	**tölee*	for
tölö-	**töle-*	to repay
ǰu:n	**ǰaun*	one hundred
ǰun	**ǰun*	summer
u:la	**aula*	mountain
ula	**ula*	sole
tu:r-	?Turkic **tü:r-*	to unwrap
tur-	**tur-*	to be hungry
mu:dən	**modun*	wood
mudən	**emüdün*	trousers

3.11.3.3. Vowel length distinctions in Mongghul

See Čenggeltei 1988, Junast 1981b. In Mongghul, most vowel length oppositions pertain to the non-high vowels.

Mongghul	source	
ala:	Amdo *ɣla*	musk deer
ala	**ala*	crotch
ula:-	**uïla-*	to cry
ula:	**ula*	sole
ula	**aula*	mountain
buda:	**budaa*	meal
buda	**bïda*	we
ʂda:-	**sïta-*	to light
ʂda-	**čïda-*	to be able
śira:-	**sïra-*	to roast
śira	**sïra*	yellow
da:li:	**dalu* (?*daluï*)	shoulderblade
dali:	**dalaï*	sea
dali:	**adalï*	alike
ta:-	**taa-*	to guess
ta	**ta*	you

bule:n	**bülien*	lukewarm[170]
bulen	**belen*	ready
ne:-	**nee-*	to open
ne	**ene*	this
ʂde:-	**ügtee-*	to weed
ʂde	**erte*	early
re:	**ere*	male
re-	**ire-*	to come
śʑi:	Amdo *hći*	middle
śʑi-	**ečі-*	to go
sgo:-	**sögee-*	to scold
sgo	**süke*	axe
ʂdo:lə-	**ötel-*	to become old
ʂdolə-	**hogtal-*	to fell

3.11.3.4. A note on the long vowels of Mangghuer

Čenggeltei has observed vowel length in a number of Mangghuer words. These long vowels are not described by Todaeva or Junast & Li. They may be the last remnants of an original system of length oppositions like that of Mongghul. Most of the recorded Mangghuer lengths can be historically explained; many stem from double or complex vowels, e.g. *gi:-* 'to release' < CM **gee-*, *mila:* 'whip' < **mïlaa*, *mo:r-* 'to low' < **möere-*; *na:du-* 'to play' < **naad-*; *o:ko* 'fat' < **öekün*; *bo:sï* 'louse' < **böesün*; *argo:sï* 'thorn' < **örgeesün*, *ġo:r* 'two' < **koar*; *qa:-* 'to cover' < **kaa-*; *sa:-* 'to milk' < **saa-*; *da:r-* 'to surpass' < **daarï-*; *do:ro* 'underside' < **doara*.

Other lengths are compensations for consonants, vowels, or syllables that were lost: *çɤ:* ~ *çɤğ* 'time' < **čag*; *no:ći-* 'to surpass, exceed' < **nögči-*; *ʂa:* 'yellow' < **sïra*. In such cases, a variant with the consonant and without the vowel length is usually also recorded. The long vowels in *be:ćin* 'illness' < **ebečin*; *be:sə* 'grass' < **ebesün* may have developed after losing the first syllable, but these forms are also reminiscent of the type of lengthening of non-high vowels followed by a high vowel, as found in Eastern Yugur and Mongghul (see 3.14.5.).

Other lengths may be accent-related, e.g. *nta:-* 'to sleep' < **unta-*, *sula:* 'loose' < **sula*; *ʑe:ni* 'one's own' < **eǰen* + genitive. Such lengths often agree with Mongghul and/or other languages, and may thus reflect regional developments of some antiquity.

[170] Thus according to Junast's materials; in Li and S/M these words have become formally identical.

3.11.3.5. A note on the long vowels of Dahejia Baoan

In Dahejia Baoan, Todaeva still observed some (rather incidental) cases of vowel length, whereas Bökh & Liu did not. It seems that the recorded lengths are relics of the more elaborate system that must still have existed in the shared ancestor of the Baoan dialects. The lengths tend to be diachronically expected, representing known contraction lengths from double or complex vowels: *bu:-* 'to descend' < CM **bau-*, *he:-* 'to warm oneself in the sun' < **hee-*, *na:tə-* 'to play' < **naad-*, *ne:-* 'to open' < **nee-*, *no:r* 'face' < **nïur*, *o:lə* 'mountain' < **aula*, *o:lon* 'cloud' < **eülen*, *su:-* 'to live; to sit' < **sau-*, *ta:-* 'to guess' < **taa-*, *to:r-* 'to wander' < **töeri-*, *tu:-* 'to drive forth' < **tau-*, *u:-* 'to drink' < **uu-*, *xa:-* 'to cover' < **kaa-*. Others represent more recent Baoan contractions, e.g. *de:sə-* 'to spread' < **debis-*, *te:rə-* 'to embrace' < *teberi-*. Not from contraction, but generally in agreement with other Baoan forms: *la:-* 'to cry' < **uïla-*, *na:-* 'to fall' < **una-*, *sa:yə-* 'to wait' < **sakï-*, *wa:r-* 'to grab' < **barï-*. All examples are from the wordlist in Todaeva 1964. These same words are documented with short vowels in other publications.

3.11.3.6. Vowel length distinctions in Ñantoq Baoan

The following examples are mainly taken from the list in Chen Naixiong 1987:46-47, which includes oppositions where both members are loanwords from Tibetan or Chinese. Again, these examples illustrate clearly that present-day oppositions do not necessarily correspond to CM oppositions. In Ñantoq Baoan, long vowels occur predominantly in the first vowel position.

BaoÑ	source	
na:raŋ	**narïn*	fine
naraŋ	**naran*	sun
ma:ra-	**marïa-*	to hide (metathetic for **mara:-* ?)
mara	**marïa*	fleshiness
sa:-	**sa:-*	to milk
sa	**basa*	also
da:rə-	**daara-*	to feel cold
darə-	**daru-*	to press
ta:-	**taa-*	to guess
ta	**ta*	you (plural/polite)
ya:r-	**yaara-*	to hurry
yar	Amdo *hyar*	summer
ke:lə	**keeli*	belly
kel-	**kele-*	to speak

teːrə-	**teberi-*	to embrace
terə	**tere*	that
ʑiː-	**ǰii-*	to extend
ʑi	**ge.ǰü*	and
ɵːləŋ	**eülen*	cloud
ɵləŋ	**olan*	many
ɵːdə	**öede*	up
ɵdə-	**od-*	to go
tɵːr(ə)-	**toarï-*	to circle
tər	**tor*	net

Note that Chen (1985) lists the forms *ɵd-*, *teːr-*, *tɵːr-* instead of the forms *ɵdə-*, *teːrə-*, *tɵːrə-* included in the list of minimal pairs in Chen (1987:46-47).

3.11.3.7. Vowel length distinctions in Xiazhuang Baoan

The examples are from Bökh & Chen, mostly from p. 62.

BaoX	source	
aːnə	?Turkic *ana*	mother
'anə	**alï* (?* *alï ïnu*)	which
χaːrə-	**kaarï-*	to scald, burn
χaːrə-	**karï-*	to return
χarə-	**gar-*	to come out
daːrə-	**daara-*	to feel cold
daːrə-	**daru-*	to press
darə-	Baoanic **šitara-*	to catch fire
doːlə-	**daula-*	to sing
doːlə-	**dolïa-*	to lick
dolə-	**hogtal-*	to cut off, sever
xoːŋ	Amdo *hoŋ(-wo*)	body
xoŋ	**hon*	year

In spite of the scarce materials it is clear that length is distinctive in this dialect. Historically, the long vowels have diverse origins, but several are regular contractions of CM double vowels, e.g. *naːtə-* ‘to play’ < **naad-*. Others are due to more recent contractions, usually involving the loss of *-w-* < **b*, e.g. *suːkə* ‘awl’ < CM **siböge*, *teːrə-* ‘to embrace’ < CM **teberi-*. The length of the *aː* in *ġaːdər*

‘sickle’ is perhaps a compensation for the loss of length in the second syllable (**kaduur*). The length of *vɛːr* ‘horn’ (**(h)eber*), without equivalents elsewhere, may compensate for the lost initial, or be due to the *-r*.

Short vowels appear instead of the expected long vowels or diphthongs in many words including: *dulə-* ‘to jump’ (**deüli-*); *kisoŋ* ‘navel’ (**köisün*); *ǰoŋ* ‘needle’ (**ǰeɯn*); *fəli-* ‘become red’ (**hulaï-*).

3.11.3.8. A note on the long vowels of Kangjia

The vowel lengths recorded for Kangjia show that its long vowels have become marginal and optional. Diachronically the remaining lengths are the usual contraction products from double or complex vowels. Examples (all from the wordlist in Sečenčogt 1999): *ba:sun ~ basun* ‘excrement’< **baasun*, *či:- ~ či-* ‘to become white’ < **čaï-*, *gi:- ~ gi-* ‘to become bright’ < **gei-*, *na:di- ~ nade-*‘to play’< **naad-*, *u: ~ u* ‘wide’< **au*, *xa:- ~ xa-* ‘to cover’ < **kaa-*, *čilə: ~ čɵlə: ~ čɵlə* ‘free time’ < **čilöe*.

3.11.3.9. A note on the long vowels of Dongxiang

Vowel lengths in native words in Dongxiang are only recorded by Todaeva (1961). Historically they are the usual contraction products. The following are all the lengths in native words listed by Todaeva: *i:n* ‘this side’ < **ein*, *qa:-* ‘to cover’ < **kaa-*, *śi:-* ‘to pull out’ < **sei-*, *xo:* ‘anger’ < **haur*, *ya:lə* ‘how’ < either **yaa-* ‘to do what’+ final converb *-lA* or **yaan* ‘what’+ comitative case **-lUA*. These same words are documented in other vocabularies without the vowel length.

3.11.4. Origins of vowel length

Historically, most vowel lengths result from the contraction of double and complex vowels, *VV* sequences which in their turn mostly stem from Proto Mongolic *VCV* sequences. Such contraction lengths are therefore traditionally called ‘secondary’ vowel lengths (Poppe, Doerfer). The contraction already took place early, and no uncontracted double vowels survive in modern languages. Many of the complex vowels do survive as diphthongs.

It is unlikely that the double vowels in Middle Mongol were anything more than spelling devices to write long vowels. Spellings such as *a’a* in Sino-Mongolic already alternated with simple *a*, which in such alternations is interpreted as *ā* by modern editors, cf. the variants *zha-ya-a-zhou* and *zha-ya-zhou* ‘destined’ in the Hua-Yi Yiyu, read *ǰaya’aǰu* and *ǰayāǰu*, respectively, by Mostaert. The two Sino-Mongolic notations are likely to represent the same pronunciation *ǰayāǰu*. Even the corresponding Written Mongol spellings (*aɣa*, etc) may be an orthographic means to indicate vowel length.

All modern languages which to some degree preserve the phonemic vowel length distinction, and tend to retain the CM double vowels as long vowels, also feature vowel lengths that are apparently not the result of contraction. Because these vowel lengths were thought, or hoped, to represent vowel lengths that predated the contraction lengths, they are traditionally called ‘primary’ long vowels (see

Janhunen 2003b for a synopsis). The lengths were viewed as possible relics from an early stage of the ancestral language, and would correspond to the long vowels of Turkic and Tungusic.

Unfortunately the 'primary' long vowels are not attested in the old sources or indicated in Written Mongol spelling. The internal Mongolic documentation for 'primary' lengths is purely modern. Evidence for non-contraction lengths is found in several central and peripheral languages.

The case for primary vowel lengths in Moghol seems weak, and will not be discussed here. Among the other peripheral languages Dagur, Eastern Yugur, Mongghul, and to a lesser extent, some Baoan dialects, also have many vowel lengths that are not the result of contractions. Unfortunately, the evidence is often contradictory, and rarely do we find vowel lengths supported by more than a single language.[171] Vowel length supported by one Mongolic language only, or by two languages in the same periphery, is insufficient evidence to assume a long vowel in CM, but may represent a regional innovation.

In fact, many of the modern non-contraction lengths can be explained as other types of secondary developments, such as more recent elisions and contractions, or compensatory lengthening. There is also a correlation between accent and vowel length, and between vowel length and vowel height, which accounts for many forms that at first sight would be potential evidence for primary vowel length. What remains is a relatively small set of unexplained lengths, which for now seems insufficient to uphold 'primary lengths' as a category.

It is important to be aware of the fact that the decision whether a Mongolic vowel length in a give lexeme is 'secondary' or 'primary' has mostly depended on a comparison with Written Mongol spellings. Secondary lengths are supposed to be the result of contraction because in Written Mongol they are spelled as *VCV* sequences (which may correspond to actual *VCV* sequences in Turkic and Tungusic cognates). However, for words that have ambiguous or uninformative non-Mongolic cognates, or lack cognates altogether, the intervocalic consonant can not be confirmed. Thus, CM **hulaan* 'red' (thus reconstructed here, but all modern forms descend from **hula:n*) is assumed to stem from trisyllabic **hulaCan* (possibly **hulaɣan*), mainly because of the Written Mongol spelling *ulaɣan*. In this approach, modern vowel lengths that are not indicated in spelling, such as that of CM **haaga* 'bran' (Written Mongol spelling *aɣa*), are candidates for 'primary' vowel lengths. But words like *aɣa* may actually just be spelling oddities without historical significance.

The various types of explicable vowel length will be discussed in separate sections below, including:

- Contraction length from double and complex vowels
- Accent-related vowel lengthening
- Compensation length

[171] The Mongolic situation can not be compared with the evidence for vowel length in Common Turkic where Turkmen, Yakut, and Khalaj, three very different languages in three different corners of the Turkic-speaking areas, largely support vowel lengths in the same words.

- Lengthening of non-high vowels preceding a high vowel (E. Yugur, Mongghul)
- Lengthening of final vowels of monosyllabic words (Dagur)
- Lengthening of the vowel of certain suffixes

I will also list examples for different types of unclear vowel length, if possible with attempts at an explanation. Perhaps some of these lengths can later be demonstrated to go back to the CM period, or perhaps their occurrence will turn out to be indirect evidence for another ancient phonetic feature.

3.12. Development of the CM double vowels

3.12.1. Introduction

In this section the usual development of the double vowels in the peripheral languages will be discussed. As indicated in chapter 2, the CM entities written here throughout as **aa*, **ee*, **ïï*, **ii*, **uu*, and **üü*, must have developed into (monosyllabic) long vowel phonemes **a:*, **e:*, **ï:*, **i:*, **u:*, and **ü:*, at the stage of development that is relevant for the discussions below. The notations used here are meant to facilitate the comparison to Written Mongol and Middle Mongol.

The complex vowels **oa*/*ua* and **öe*/*üe* will also be discussed here, as at the same stage of CM they can be assumed to have developed into long vowels **o:* and **ö:*. With regard to these two vowels, the digraph notations are again a concession, in this case also to Moghol, which seems to preserve these sequences as diphthongs.[172]

The modern developments of **a:*, **e:*, **ï:*, **i:*, **o:*, **ö:*, **u:*, and **ü:* tend to be more predictable than their short counterparts, which are prone to changes effected by the phonetic environment, reduction and elision. However, even in those modern languages that have not given up the phonemic length distinction, the double vowels can be contracted and shortened, and then may undergo any development found in the short vowels, including complete loss, as in BaoÑ *hda* < **huuta* 'bag'.

In addition, many of the complex vowels have been simplified and merged with the long vowels, cf. MgrH *ra:* < CM **arïa* 'molar tooth'.

All this leads to many possible origins for each modern phoneme, and conversely, many modern reflexes for every CM phoneme. The long vowel *o:* in Mongghul illustrates this problem:

bo:-	**boa-* (**bo:-*)	to bundle
xo:ʒən	**koasun* (**ko:sun*)	empty
bo:sə	**böesün* (**bö:sün*)	louse
ʑo:lon	**ǰöelen* (**ǰö:len*)	soft
o:sgu	**auškï*	lungs
ġo:r	**koar*	two

[172] At least in Ramstedt's materials. The question whether the Moghol diphthongs could be a secondary development will not be discussed here.

(continued)

no:r	**noïr*	sleep
fo:r	**heür*	nest
bo:lə	**büile*	gums
noġo:n	**nogaan*	green
sgo:-	**sögee-*	to scold
furo:-	**horïa-*	to wind
fo:de	**hodun*	star
fo:de	**hödün*	feather
śilo:	**silön*	soup
muro:n	**mören*	river
mo:r	**mör*	weg
o:sə-	**ös-* and/or **eüs-*	to grow
yo:-	**oya-*	to sew

In the following the general tendencies will be given for the development of each double vowel; the focus will be on the languages preserving phonemic vowel length.

3.12.2. Development of double **aa*

Distribution

The sequence **aa* can be found in all vowel slots:
Word-initially: **aalï, *aasï* 'character' (not listed), **aarča* 'curds';
Non-initially in the first vowel slot: **baa-* 'to defecate', **čaalsun* 'paper', **daagan* 'foal', **daarï* 'saddle gall', **kaa-* 'to close', **maajï-* 'to scratch', **naad-* 'to play', **saa-* 'to milk', **taa-* 'to guess', **yaara-* 'to hurry';
In non-first vowel slots: **budaan* 'porridge', **čagaan* 'white', **dutaa-* 'to flee', **hulaan* 'red', **hutaa* 'smoke', **ïmaan* 'goat', **ǰuǰaan* 'thick', **kïǰaar* 'edge', **mïnaa* ~ **mïlaa* 'whip', **salaa* 'branch', **umdaas-* 'to be thirsty'.

Developments

Double **aa* is usually preserved as *a:* in all languages which preserve the feature of vowel length. In Ñantoq Baoan *a:* is usually shortened beyond the first syllable. In Mangghuer, Dahejia Baoan, Dongxiang and Moghol, long *a:* < **aa* has merged with its simple counterpart *a* < **a*. In the languages preserving length, *a:* does not undergo most of the pronunciation changes (reduction, palatalisation, labialisation) which occurred in **a*.

CM **aa* of the second syllable remains **a*: after **a, *ï,* or **u* of the first syllable. Examples:

CM	Dag	EYu	MgrH	BaoÑ	
**baa-*	*ba:-*	*pa:-*	*ba:-*	*ba:-*	to poo
**naad-*	*na:d-*	*na:d-*	*na:də-*	*na:d-/na:tə-*	to play
**ǰalaa*	*ǰila:*	*ǰala:*	*ʒ́ala:*	*ʒ́ala*	tassel
**čagaan*	*čiɣa:n*	*čağa:n*	*ćaġa:n*	*ćiχaŋ*	white
**ïmaan*	*ima:*	*ma:n*	*ima:*	*yimaŋ*	goat
**hulaan*	*xula:n*	*ła:n*	*fula:n*	*fulaŋ*	red

There seem to be no certain examples for the existence in CM of a vowel sequence **o-*aa*. Poppe (1955:60) does not list examples for the sequence *CoCaɣa*. There is no reason why it could not occur, but evidence from the modern languages does not allow us to distinguish between **nogaan*, **noguan*, or **nogoan*, as these forms would result in the same modern forms. Only in some words there are historical arguments for a rounded vowel in the second vowel slot. For instance, in the case of **doloan* 'seven' there must once have been a rounded element in the second syllable to explain the **o* of the first syllable, in view of its relationship with **dalan* 'seventy'. Otherwise the sequence **oa* can be distinguished from **aa* only after a first syllable with **ï* (for **oa* see below). As a pragmatic solution the notation with **aa* is chosen here, e.g. **nogaan*. In most words listed below reconstructions with **aa*, **ua*, or **oa* would all be equally defensible.

CM	Dag	EYu	MgrH	BaoÑ	
**boraan*	---	*bɔrɔ:n*	*buro:n*	---	rain
**dolaan*	*dɔlɔ:(n)*	*dɔlɔ:n*	*dolo:n*	*dɵlɵŋ*	seven
**nogaan*	*nuwa:*	*nɔğɔ:n*	*noġo:*	*nɵġuŋ*	green
**soyaa*	*sɔyɔ:*	*sɔyɔ:*	---	---	stalk
**togaan*	*tuwa:*	*tɔğɔ:n*	*toġo:*	*tɵχɵŋ*	pot

The sequence **ua* as such does not seem to occur in monomorphemic stems. When the sequence **ua* arises in derivation or inflection, its development in the peripheral languages can not be distinguished from that of **aa*, except when the first syllable vowel is **ï*. The same applies to most central languages.[173] Examples include **ab-* 'to take' > **ab.u.ad* 'having taken', **dusu-* 'to drip' > **dusu.a-* 'to cause to drip', **unu-* 'to ride' > **unu.a* 'saddle animal', **hutu-* 'to emit smoke' > **hutu.a* 'smoke', **ǰïru-* 'to paint' > **ǰïru.ačï* 'painter'.

Many CM words with the complex vowel **ïa* in the second syllable have developed similar modern forms as **aa* in the QG languages, especially in Eastern Yugur and Baoan (see 3.13.6.).

[173] In Khamnigan it is possible to distinguish **aa* of non-first syllables from **ua*, even if the first vowel is not **ï*, as in *abo:d* 'having taken', from **ab-* 'to take' with connective vowel and perfect converb ending *+Ad* (see Janhunen 2003:24). In other languages forms derived from an assimilated *aba:d* < **ab-u-ad* occur. Ordos has *awa:t*, however if the *u* belongs to the stem it is preserved, e.g. *alχ^{u}a:t* from **alku-*.

Even in the QG languages that preserve vowel length, *a:* is prone to be shortened. In Ñantoq Baoan this is more or less regular. In Eastern Yugur and Mongghul it happens less predictably, sometimes in the same stems, e.g. Eastern Yugur and Mongghul *sala* 'branch' < **salaa*.

3.12.3. Development of double **ee*

Distribution

The sequence **ee* can be found in all vowel slots:
Word-initially: **eede-* 'to coagulate' (not listed), **eemeg* 'earring' (not listed);
Non-initially in the first vowel slot: **čeejǐ* 'chest', **deel* 'garment', **deere* 'above', **deesün* 'rope', **gee-* 'to release', **hee-* 'to dry in the sun', **keeli* 'belly', **nee-* 'to open', **teermen* 'mill';
In non-first vowel slots: **emeel* 'saddle', **gegeen* 'bright', **güǰeen* 'stomach', **herbeekei* 'butterfly', **idee* 'pus', **siree* 'table', **sübee* 'flank', **temeen* 'camel', **ügtee-* 'to weed'.

Developments

Double **ee* typically appears as *ə:* in Dagur, but developed into a closed front vowel in QG. In Eastern Yugur, long *e:* has often been raised > *i:*, as in EYu *de:re ~ di:re* 'above' (**deere*), although Qinglong dialect preserves *e:* in these cases (see Bolčuluu & Jalsan 1988:351).

CM	Dag	EYu	MgrH	BaoÑ	
**nee-*	*nə:-*	*ni:-*	*ne:-*	*ne:-*	to open
**ǰee*	*ǰə:*	*ǰi:, ǰe:*	*ʒe:*	---	cousin
**keeli*	*kə:lʲ*	---	*ke:le*	*ke:lə*	belly
**gegeen*	*gəɣə:n*	*gegen*	*gəge:n*	*gegaŋ*	bright
**güǰeen*	*guǰə:*	*gu̵ǰe:n*	*guʒe:*	---	stomach

Rounding of **ee* after **ö* is applied less strictly than in the case of short **e* after **ö*, and **aa* after **o*.

CM	Dag	EYu	EYuQ	BaoÑ	
**nögee*	*nuwə:*	---	---	*nəgə*	other
**döree*	[*du:rə:ŋgi:*]	*du̵re:*	*döre:*	---	stirrup
**göreesün*	*gurə:s*	*gu̵re:sən*	*görö:sən*	---	wild animal
**sögee-*	*suwə:-*	---	---	*su̵gu̵-* (Kgj)	to scold

Many instances of **ie* and **üe* developed like **ee*. Like **ua* discussed above, the sequence **üe* does not occur in monomorphemic stems, but it can arise when suffixes are added, as in **bürkü-* 'to cover'> **bürkü.esün* 'covering', **hürgü-* 'to be frightened' > **hürgü.e-* 'to frighten', **mörgü-* 'to bow' > **mörgü.ed* 'having

bowed'. In all languages except Khamnigan **üe* has become indistinguishable from **ee* except when the first syllable vowel is **i*. See the complex vowels below.

Cases of shortening include: EYu *emel*, MgrH *imel* < **emeel* 'saddle', EYu *šere* < **siree* 'table', EYu *termen*, MgrH *tərme*, BaoÑ *termaŋ* < **teermen* 'mill', EYu *kʉte-* < **ügtee-* 'to pull up weeds'. In EYu *örlö ~ örlö:* < **öreele* 'half; hobble' the original long **ee* was lost altogether.

3.12.4. Development of double **ïï* and **ii*

These two double vowels, the rare double counterparts of **ï* and **i* are hard to distinguish in most languages. The difference between the short counterparts was almost annihilated in the first syllable in central Mongolic due to palatal breaking (see 3.6. above), so that they merged into one /i/ phoneme which is largely restricted to non-first syllables. Double **ïï* and **ii* are not subject to breaking, and in that sense are more stable. However, the number of examples is small.

Distribution

The relatively rare sequences **ïï* and **ii* are can be found in non-initial vowel slots:
Non-initially in the first vowel slot: **čïïrag* 'firm', **čiig* 'moist', **čiirsün* 'mat' (not listed), **ǰii-* 'to extend',**nii-* (?**nïï-*) 'to blow the nose', **sïïra* 'leg';
In non-first vowel slots: **hanïï-* (~ **hanï-*) 'to close one's eyes', ?**ǰihiin* 'purple', **ǰimii-* 'to close the mouth' (the latter two not listed).

Developments

Both **ïï* and **ii* normally developed into long *i:*. In the central languages Khalkha and Kalmuck, stems with **ïï* like **čïïrag* 'strong' and **sïïra* 'leg' were transferred to the front vowel class, even when they contained back vowels or uvular allophones of **k*, **g*, or took back suffixes. However, Ordos, Buriat and Khamnigan tend to retain the original back features. Also the QG languages allow stems in which **a* follows *i* or *i:* of the first syllable.

CM	Dag	EYu	MgrH	BaoÑ	
**ǰii-*	*ǰi:-*	---	*ʒ́i:-*	*ʒ́i:-*	to stretch
**ǰïlïï-*	---	*ǰili:-*	---	---	to be separated[174]
nii-* (?nïï-*)	*ñi:-*	---	---	---	to blow one's nose
**čïïrag*	---	*čiraġ*	*ćiraġ*	---	strong, tight[175]
**sïïra*	---	*ši(:)ra*	*śira:*	---	leg[176]
**čiig*	---	*či(:)g*	---	---	moisture[177]

[174] This word is reconstructed with back vowels because of the EYu caus. *ǰili:lġa-*.
[175] In Dongxiang we find the unexpected form *ɕiğara* with intervocalic consonant. See 4.5.4.
[176] Another unexpected Dongxiang form with intervocalic consonant: *şiğara*.
[177] Not **čïïg,* in spite of its back-vocalic Turkic cognate **či:g*.

In Eastern Yugur long *i:* can be shortened without merging with its short counterpart /i/, as the latter is usually realised as [*ə*].

3.12.5. Development of double **uu*

Distribution

The sequence **uu* can be found in all vowel slots:
Word-initially: **uu-* 'to drink', **uuča* 'back',**uurag* (?< **aurag*) 'biestings';
Non-initially in the first vowel slot: **buudaï* 'wheat', **huuta* 'bag', **ǰuura-* 'to knead', **suu* 'armpit';
In non-first vowel slots: **čïlbuur* 'martingale', **huruu* 'downward', **huruul* 'lip', **kuruun* 'finger', **nïruun* 'spine', **turuun* 'hoof'.

Developments

In Mongghul the long rounded high vowels of the second syllable are often reduced, especially words ending in this vowel (or that lost final **-n* as is common in Mongghul). This reduction is uncommon in Eastern Yugur. In Ñantoq Baoan all long vowels tend to be shortened beyond the first syllable.

CM	Dag	EYu	MgrH	BaoÑ	
**uu-*	*ɔ:-*	*u:-*	*u:-*	*u:-*	to drink
**huruu*	*xɔrɔ:*	*hɔrui, hurü:*	*furə*	---	downwards[178]
**nïruun*	*nirɔ:*	*nuru:n*	*nurə*	*nərəŋ*	back, ridge
**kuruun*	*xɔrɔ:*	*χuru:n*	*xurə*	*ġurə*	finger

In Dagur **uu* can be broken (like its short counterpart) when a short **a* followed in the next syllable[179]; the **uu* is reduced in the process, and its length transferred to the imported *a*. In Eastern Yugur, Mongghul and Ñantoq Baoan there are several instances of shortening or elision of **uu*.

CM	Dag	EYu	MgrH	BaoÑ	
**uurag*	*wa:rəl*	*uraġ*	*uraġ*	*orġə* (BaoGt)	biestings
**huuta*	*x^{w}a:t*	*u:ta*	*fu:da*	*hda*	sack
**ǰuura-*	*ǰ^{w}a:r-*	*ǰü(:)ra-*	*ʒ́ira:-*	*ʒ́ira-*	to mix

Many CM words with the complex vowel **au* have developed similar modern forms in QG (and elsewhere). The same applies to the complex vowel **ïu*. See the section on the complex vowels below.

---

[178] The Eastern Yugur form is irregular in that it seems to include a diphthong with a palatal element (also reflected in Kalmuck *ürü:*).

[179] This can also be observed in words with unexpected lengths, such as Dag *x^{w}a:l* < **huula* 'tinder'and *t^{w}a:ləɣ* < **tulga* 'support' discussed in 3.14.2.

3.12.6. Development of double **üü*

Distribution

The sequence **üü* can be found in all vowel slots:
Word-initially: **üüče* 'meat provisions for winter', **üür-* 'to carry on the back' (these words not listed);
Non-initially in the first vowel slot: **büürge* 'pommel', **düüre-* 'to fill', **tüükei* 'raw';
In non-first vowel slots: **bidüün* 'coarse', **hüǰüür* 'tip', **küǰüün* 'neck', **sirüün* 'coarse'.

Developments

Parallel to **u:*, in Mongghul **ü:* tends to be reduced at the end of words.

CM	Dag	EYu	MgrH	BaoÑ	
**hüü-*	*xu:-*	*hü:-*	*fu:-*	*fu-*	to rot
**tüükü*[180]	*tuiyun*	*tü:kə:, tü:^{h}gə*	*tiu:gu*	*tugu*	raw
**bidüün*	*budu:n*	*bodü:n*	*budən*	*bedəŋ*	coarse

(continued)

CM	Dag	EYu	MgrH	BaoÑ	
**küǰüün*	*xuǰu:*	*gʉǰü:n*	*guʒə*	*guʑuŋ*	neck
**küün*	*xu:*[181]	*kʉn, ku:n*	*kun*	*kuŋ*	person

Many CM words with the complex vowel **eü* have developed similar modern forms in QG (and elsewhere). Depending on the language, **iü, *üi and *uï* may also develop similarly. See the section on the complex vowels below.

In Eastern Yugur the long *ü:* is frequently shortened, as in *kü:ken ~ küken* 'boy' < **keü.ken* 'child'. This does not cause phonological problems, as the palatal pronunciation prevents it from merging with *ʉ*, the usual reflex of short **ü*.

3.12.7. Development of **oa* (**oo*)

Distribution

The sequence **oa* can be found in all vowel slots:
Word-initially: *oačï-* 'to drink';
Non-initially in the first vowel slot: **doa-ra* 'underside', **koa-* 'to dry', **toan* 'number', **toasun* 'dust';

---

[180] Rather than the central Mongolic form **tüükei*. Dag *-ui-* and *-n* are unexpected.

[181] Dagur also has shortened forms of **küün* in the compounds *əmyun* 'wife, woman, female' from**eme küün*, and *əryun* 'husband, man, male', from **ere küün.*

In non-first vowel slots the presence of **oa* can only be established after**ï*: **hïroar* ~ **hïraur* 'bottom', **jïloa* 'rein'. Elsewhere it has become indistinguishable from **aa*.

Developments

In all modern languages except Moghol, CM **oa* has been contracted into *o:*. In the first syllable this was preserved as a long *ɔ:* sound in Dagur unless there was an **a* in the next syllable (which triggers labial breaking). In Eastern Yugur the contracted vowel is often raised to *u:*.[182] CM **oa* merged with its harmonic counterpart **öe* in Mongghul and Ñantoq Baoan, but stayed long.

CM	Dag	EYu	MgrH	BaoÑ	
**oačï-*	---	*u:čə-*	*ućə-*	*ać**a**l-* (caus.)	to drink
**koa-*	*x^{w}a:-*	*χu:-*	*χo:-*	*χɵ:-*	to dry
**doara*	*d^{w}ar*	*du:ra*	*do:ro*	*dɵ:rə*	underside
**toala-*	*t^{w}a:l-*	*tu:la-*	*to:la-*	*tɵ:la-*	to count
**čoara-*	---	*ču:ra-*	*ćo:ra:-*	*ćɵ:ra-*	to be pierced
**toarï-*	*tɔ:r^{y}-*	*tɔrɔ-*, *toro:-*	*to:rə-*	---	to circle

In Mangghuer and Dahejia Baoan the typical reflex is short *o*, which in Kangjia may develop into *u*. In Dongxiang the situation is more complicated. In some of the stems listed above, Dgx also has *o*, e.g. *oɕï-* 'to drink', *ġo-* 'to dry'. In other stems Dgx features the diphthong *əu*, as in *dəura* 'underside', *təula-* 'to count', *ɕəura-* 'to be pierced'. These diphthongs may be archaisms. They are reminiscent of Moghol, where we find *qua-* 'to dry', *toala-* 'to count', etc, alongside *dora* 'underside', *qo:lɛı* < **koalaï* 'throat'. Moghol was one of the reasons why the CM forms are reconstructed here with **oa* rather than **o:*. However, it should be noted that Dongxiang and Moghol do not generally have the diphthongs in the same stems, and Dongxiang *əu* has several, partly unclear origins.[183]

The broken vowels in Dagur are not instances of preserved uncontracted **oa*, but are due to labial breaking due to the presence of **a* in the next syllable (cf. Dag *tɔ:(n)* 'number' from **toan* as opposed to its derivation **toala-* above).

The developments in the second vowel slot are similar.

CM	Dag	EYu	MgrH	Kgj	
**hïroar*	*xiɣɔ:r*	*hru:r*	---	*hurlu*	bottom
**jïloa*	*ǰilɔ: ~dilɔ:*	*ǰəlu: ~ ǰulu:*	*ʒolo.ćə*	---	reins
**jïroa*	*ǰirɔ:*	*ǰɔrɔ:*	*ʒoro*	*ǰuru*	ambler

[182] Qinglong dialect tends to preserve *ɔ:*. See Bolčuluu & Jalsan (1988:350-351) for dialectal differences regarding this. However, as in the case of **öe* discussed below, forms such as *ǰɔ:n* 'hundred', *xɔ:čən* 'old' (from **ǰaun, *kaučïn*) look like results of secondary lowering.

[183] Among other things, it occurs in stems where a final *-g* was lost, e.g. *odəu* 'knee' from **ebüdüg*.

3.12.8. Development of **öe* (**öö*)

Distribution

The sequence **öe* can be found in all vowel slots:
Word-initially: **öede* 'up', **öer-* 'self', **öere* 'other'.
Non-initially in the first vowel slot: **böelǰi-* 'to vomit', **böere* 'kidney', **böesün* 'louse', **höesün* 'pus', **köesün* 'scum', **köerge* 'bridge', **töeri-* 'to get lost'.
In non-first vowel slots the presence of **öe* can only be established after**i*: **čilöe* 'free time', **hiröe-* 'to bless', **kiröe* 'saw'.

Developments

The developments are parallel to those of **oa*. In all modern languages except Moghol, CM **öe* has been contracted. Dagur has the developments *u:* and *wə:*. The latter is not a continuation of **öe*, but due to labial breaking (probably partly triggered by **e* of the second syllable). The different treatment of Dag *bə:r* (for **b^{w}ə:r*) < **böere* 'kidneys', and *bu:s* < **böesün* 'louse' is probably due to the different vowels of the second syllable.[184] In Eastern Yugur it stays long, but is often raised to *ü:*.[185] In Mongghul and Ñantoq Baoan, **öe* merged with its harmonic counterpart **oa*, and resulted in long *o:*, which was shortened in Mangghuer and most Baoan dialects.

CM	Dag	EYu	MgrH	BaoÑ	
**köe-*	*x^{w}ə:-*	*kü:-*	*ko:-*	*kɵ:-*	to swell
**čöen*	*č^{w}ə:n*	*čü:n, čö:n*	*ćo:n*	*čoŋ* (BaoX)	few
**böere*	*bə:r*	*pü:re, pö:rö*	*bo:ro*	*bɵ:rə*	kidneys
**böesün*	*bu:s*	*bü:sən*	*bo:sə*	*bɵ:səŋ*	louse
**böelǰi-*	*bə:lǰ-*	*bölǰö-*	*bo:lʑo-*	*bɵlʑa-*	to vomit

In Dongxiang the development of **öe* differs from that of **oa*, as **öe* has never resulted in the diphthong *əu*.

In Dagur palatal breaking is less common than in the back-vocalic stems above. The second vowel is unrounded, whether or not breaking took place. The quality of the second syllable in QG is erratic.

---

[184] Likewise the Dagur variants of the same etymon *t^{w}ə:r-* and *tu:r^{y}-* 'to get lost' perhaps reflect CM variants **töere-* and **töeri-*.

[185] Qinglong has *ö:*. See Bolčuluu & Jalsan 1988: 350-351 for examples for this dialectal difference. However, some Qinglong forms raise the question whether some cases of *ö:* are not the result of secondary lowering. This seems to be the case for *dö:* 'younger brother', *dölö:n* 'spleen', *törö:* 'bootleg' (from **deü, *deliün, *türei*).

CM	Dag	EYu	MgrH	BaoÑ	
**čilöe*	*čulə:*	*čölö:*	---	*ćələ*	free time
**hiröe-*	*xirə:-*	*hörö:r*	*śüro:-*	---	to bless
**kiröe*	*kirə:*	*kʉre:*	*ćirəu:*	*kuru*	saw

MgrH *śʒo:sə* suggests a reconstruction **hičöesün* ‘willow’ (as does Buriat), but EYu *hərči:sən* and Kgj *jasun* favour **hičesün*.

3.13. Development of the complex vowels

3.13.1. Types and developments

The complex vowels, or diphthongs, are sequences of two different vowels. They are called complex vowels here, because it is not known whether they originally constitute one or two syllables. There are some indications that at least in some words they were originally disyllabic.

Firstly, vowel sequences may span two morphemes, e.g. **tarïan* ‘crop’ < **tarï-* ‘to plant, cultivate’, **kalaun* ‘hot’ < **kala-* ‘to be hot’.

Secondly, the elements of a complex vowel may behave independently in that they can be seen to swap places. For this type of metathesis see 3.13.6.

Thirdly, the diphthongs in some words are known, mostly on the basis of non-Mongolic data, to go back to VCV sequences. Examples include **aïl* ‘campsite’ (Turkic **agïl*) and **siür-* ‘to sweep’ (Turkic **sipür-*).

The historical origin and morphological structure of the complex vowels has no bearing on their further development in the peripheral languages, that is, it is irrelevant whether there is evidence for a lost intervocalic consonant.

All complex vowels are made up of vowels that also occur independently. In many respects the diphthong elements behave like their independent counterparts. For example the **a* of **aï* may be harmonically rounded by a preceding **o*, and the **i* of **ie* palatalises a preceding **s*. In other repects they are different entities, for instance the second element of **aï*, **oï*, **uï*, retains its palatality in the QG languages, whereas **ï* and **i* generally become centralised into a *ə*-like sound.

For the purpose of discussing their development in the QG languages, the CM complex vowels can be divided in four groups.

1) Diphthongs ending in **ï/*i* **aï*, **ei*, **oï*, **öi*, **uï*, **üi*

2) Diphthongs ending in **u/*ü* **au*, **eü*

3) Diphthongs starting with **ï/*i* **ïa*, **ie*, **ïu*, **iü*

4) Diphthongs starting with **u/*ü* **ua*, **üe*

The first three groups survive in Dagur, Moghol, and in both branches of Shirongol, although there are many contractions and exceptions in individual Shirongol

languages. In Eastern Yugur only group 1 partly survives.[186] Eastern Yugur thus even exceeds Khalkha, Ordos, and Kalmuck, which incidentally preserve group 3 diphthongs in non-first vowel slots. Within Shirongol, Ñantoq Baoan is perhaps the most eagerly contracting dialect, with groups 2-4 always resulting in monophthongs, and all of group 1 often appearing as *i:*.

The diphthongs **ua/*üe* of group 4 were assimilated to either *aa/ee* (after *A* or *U* of the first syllable) or *oo*/*öö* (in the first vowel slot or after **ï/*i* of the first syllable).[187] Therefore they are discussed in the section on the long vowels above. They are written **oa* and **öe* in the reconstructions in order to facilitate comparison with written forms of Mongolic and with Moghol, the only modern language that did not systematically monophthongise these two sequences. This notation will also facilitate the comparison with non-Mongolic data.

Partly due to the simplification of vowel harmony, not all of the original diphthongs survive in native words in the QG languages.

Phonetic changes that affect the diphthongs may vary greatly, even within languages. The main changes are partial assimilation or complete assimilation (monophthongisation). The resulting monophthongs may subsequently be shortened, or, especially in the Shirongol languages, diphthongised into new (Chinese-inspired) diphthongs. These predictable diphthongal pronunciations will not be indicated in the notations here.

The second diphthong element tends to be dominant. Group 1 diphthongs ending in **ï/*i* will always retain a palatal element, whereas the preceding element may be partially or fully assimilated. Group 2 diphthongs ending in **u/*ü* will always retain a labial element. In group 3 the palatal first element may be assimilated to the second element.

The elements *a* and *e* of the diphthongs **aï*, **ei* and **ïa*, **ie* are subject to harmonic rounding when following **o/*ö* of the first syllable, as in EYu *mɔǧɔi* < **mogaï* 'snake'.

The merger and loss of vowel sequences in Shirongol can be partially predicted from the development of the simple vowels, e.g. the loss of the distinction between **uï* and **üi* follows from the loss of the distinction between **u* and **ü*. Diphthongs absent from Chinese may shift towards a Chinese-compatible pronunciation, e.g. **oï* > *uai* or *ai* or *ei*.

A noteworthy special development in Dagur is that diphthongs can undergo palatal breaking, as in $x^{y}aut$ < **hïutan* 'narrow', as well as labial breaking, as in $x^{w}ainə$ < **koïna* 'posterior'.

As similar developments have taken place repeatedly, in different periods (already starting in Middle Mongol), and in different branches, diphthong developments are of limited use for classification purposes.

In the following discussion it has to be kept in mind that many of the complex vowels that had been assimilated in a given language may re-emerge at a later stage, due to vowel reduction in native words. For instance the sequence **ïa*

---

[186] In the central languages the treatment of the vowel sequences varies greatly. In Buriat most diphthongs of groups 1 and 3 are preserved. In Chakhar group 2 and most of group 1 have become monophthongs; the same has often happened to the sequences of group 3.

[187] Only Khamnigan shows a different development, e.g. *abo:d* < **ab.u.ad* 'having taken'.

had been assimilated to *a:* in earlier Eastern Yugur, but modern Eastern Yugur has developed a new diphthong *ia* which is the result of the reduction of an unstressed vowel, as in *biar* < **bayar* 'joy'. The *-u-* in MgrM *ŋġua-* < **ugaa-* < **ugïa-* 'to wash' echoes the preceding rounded vowel. The status of such sequences may be reinforced by Chinese loanwords containing similar diphthongs.

3.13.2. Development of **aï* and **ei*

3.13.2.1. Distribution

The sequence **aï* can be found in all vowel slots:
Word-initially: **aïl* 'village', **aïmag* 'village';
Non-initially in the first vowel slot: **baï-* 'to stand', **čaï-* 'to dawn', **kaïčï* 'scissors', **kaïla-* 'to call', **naïman* 'eight', **naïta-* 'to sneeze', **saïkan* 'beautiful', **taïl-* 'to untie';
In non-first vowel slots: **čïraï* 'face', **gakaï* 'pig', **horaï* 'crown', **hulaï-* 'to become red', **mogaï* 'snake', **nokaï* 'dog', **taŋlaï* 'palate', **taulaï* 'hare'.

The sequence **ei* can be found in all vowel slots:
Word-initially: **ein* 'this way';
Non-initially in the first vowel slot: **gei-* 'to become bright', **geičin* 'guest', **heil-* 'to leave', **kei* 'wind', **neile-* 'to unite';
In non-first vowel slots: **dülei* 'deaf', **herbeekei* 'butterfly', **hümekei* 'smelly', **isegei* 'felt', **melekei* 'frog', **silükei* 'saliva'.

The occurrences in non-first syllables are all word-final.

3.13.2.2. Dagur metanalysis of final **aï* and **ei* in noun stems

In nouns ending in **aï* and **ei*, these diphthongs tend to lose their second element, and as a consequence may disappear altogether. However, Dagur adjectives and verbs generally preserve the diphthongs as such, or monophthongised into a long *i:*. This division is strikingly illustrated by the contrast between the CM noun **kauraï* 'file (tool)', which developed into Dag *xaur*, whereas the homophonous adjective **kauraï* 'dry' developed into Dag *x*^{w}*a:ri:*.[188] This indicates that the loss of the second diphthong element is not a phonetic matter. Rather, it seems to be due to metanalysis, whereby the bare stem was 'mistaken' for its genitive-accusative, and a new stem was created by removing the perceived case ending. Examples:

Dagur	CM	
gaɣ(ə), gen-acc *gaɣəi*	**gakaï*	pig
nɔɣ^{w}, gen-acc *nɔɣui*	**nokaï*	dog
mələɣ, gen-acc *məlɣi:*	**melekei*	frog

[188] Other adjectives (all preserving a diphthong or contracted long vowel in Dagur) include **kelegei* 'mute', **köŋdei* 'hollow', **solagaï* 'left', **subaï* 'barren'.

This reinterpretation has occurred in many (most?) noun stems, but is not universal, cf. Dag *dali:* < **dalaï* 'sea'.

Another clue as to why Dagur lost the **-ï/-i* from word-final diphthongs may be provided by Manchu cognates (in most cases obvious borrowings from Mongolic) which also end in a simple vowel rather than a diphthong. Examples:

Dagur	as if from CM	cf. Manchu	
gaɣə	**gaka* (for **gakaï*)	---	pig
ǰɔl	**ǰula* (for **ǰulaï*)	*giyolo*	fontanel
maɣəl	**magala* (for **malagaï*)	*mahala*	hat
mɔɣ^{w}	**moga* (for **mogaï*)	---	snake
šar	**čïra* (for **čïraï*)	*cira*	face
tannə	**taŋla* (for **taŋlaï*)	---	palate
xaur	**kaura* (for **kauraï*)	*hûwara*	file
x^{w}*alaɣ*	**kulaga* (for **kulagaï*)	*hûlha*	thief
x^{w}*arəm*	**korma* (for **kormaï*)	---	hem
xɔr	**hora* (for **horaï*)	*foron*	crown of the head
xɔruɣ^{w}	**koraka* (for **korakaï*)	---	insect

Similarly, in front-vocalic nouns:

kumuɣ	**kömüge* (for **kömügei*)	---	uvula
mələɣ	**meleke* (for **melekei*)	---	frog
urum	**ölme* (for **ölmei*)	---	instep
tur	**türe* (for **türei*)	*ture*	bootleg
xərəɣ	**hereke* (for **herekei*)	*ferhe*	thumb

There are two possible explanations for the similarity of the Dagur and Manchu forms. Dagur could have taken these word shapes from Manchu, although several of them are not actually attested in Manchu. At any rate the Manchu forms need explanation themselves. They could stem from an Oirat-type language that monophthongised **aï* and **ei* into *a:* and *ä:* respectively. Alternatively, the Manchu forms are reminiscent of the Dagur ones because Manchu took many of its Mongolic loanwords from Dagur, as suggested by Doerfer (1984:11-12). In that case metanalysis seems the most likely reason behind the Dagur forms, at least provisionally.

Metanalysis may be at the root of similar cases in QG languages, where the loss of the last diphthong element occurs sporadically, e.g. EYu *ölöɣö* < **ölegei* 'cradle', MgrH *śolgo* < **silükei* 'saliva', MgrH *malġa* (with equivalents in the other Shirongol languages) < **malagaï* 'hat', MgrM *qormo* < **kormaï* 'hem'.[189] However, among the QG languages, only Eastern Yugur preserves the suffix shape *-i:n* as one of the variants of the genitive-accusative case ending; this suffix variant is required for metanalysis to take place. In Shirongol the shape *-nI* has become generalised.

[189] **kïtuga(ï)* 'knife' only preserves the element **ï* in Moghol.

Any metanalysis in Shirongol based on reinterpretation of the nominative as a genitive must have taken place in an earlier period.

On the other hand old variants found throughout Mongolic, such as **andagar ~ andagaï* 'oath', **ǰulaï ~ *ǰula* 'fontanel', **koalaï ~ *koala* 'throat; food', may indicate that the **-ï/-i* is an obsolete formans of unclear function. This is also suggested by the fact that words may lose the element **ï/*i* before some suffixes, cf. the plurals of **nokaï* 'dog', in SH *noqo-t*, Kalmuck *noχǒ-s*, 'dogs'. Whatever the original etymological status of the element **ï/*i* , a singular without it may have developed due to reanalysis of the plural.

3.13.2.3. Remaining developments of **aï* and **ei*

Initially and in monosyllables **aï* is generally preserved as *ai* both in Dagur and the QG languages. In non-first vowel slots **aï* may appear in several degrees of assimilation and contraction, i.e., *ai, əi, ei, e:, i:,* which may then be shortened to *e, i*. The reflexes often vary within a language, e.g. Dag *dalai ~ dali:* < **dalaï* 'sea', EYu *aŋğai- ~ aŋğəi- ~ aŋği:-* < **aŋgaï-* 'to open', MgrH *će:- ~ ći:-* < **čaï-* 'to become bright'. Likewise **ei* may be assimilated, as in Kgj *sei- ~ si-* < **sei-* 'to weed'. In Mangghuer and Dongxiang, **ei* may appear as *ai***,** as in Dgx *lailie-* < **neile-* 'to unite'.

As discussed above many nouns ending in **aï* and **ei* lose the second element of these diphthongs in Dagur, so that adjectives and verbs are the most reliable examples. After a rounded vowel, or partly rounded diphthong, *-*aï* will result in Dag *-ui*.[190]

Dag	CM	
aulɣui	**aulïkaï*	fearful
ñɔskui	**nïsukaï*	snotty
sɔɣui	**subaï*	barren
sɔlɣui	**solagaï*	left-handed

Rounding of **aï* after *ɔ* is also found in Eastern Yugur and Shirongol, e.g. EYu *sɔlɔğɔi* < **solagaï* 'left-handed', BaoD *noġoi* < **nokaï* 'dog'. This rounding is not present in Dongxiang. It was probably secondarily lost, as the sequence *o-o* is allowed in Dongxiang.

[190] In some Dagur words **-aï* behaves like the complex vowel **ïa,* and develops into prepalatalising *e:,* as in *ɔmʸe:* < **umaï* 'Umay (goddess of childbirth)', *ɔrʸe:* < **oraï* 'evening', *ñerʸe:* < **nïraï* 'tender', *če:-* < **čaï-* 'to become white'. In the last example the palatal consonant may have played a role. Another group of words reduced the original **aï/*ei* into a palatalising element after the preceding consonant, e.g. *bə:lʸ* < **beelei* 'gloves', *čakɔ:lʸ* < **čakulaï* 'gull', *taulʸ* < **taulaï* 'hare', but this may merely be the same phenomenon that caused the palatalisation of word-final *-l* and *-r*, as in *kulʸ* < **köl* 'foot' and *gərʸ* < **ger* 'house'. Yet another strange development, reminiscent of Oirat, is Dag *jauǰa:* < **ǰauǰaï* 'heel'.

3.13.3. Development of **oï, *uï *öi, *üi*

3.13.3.1. Distribution

Compared to **aï* and **ei*, the sequences **oï, *uï, *öi, *üi* were rare, especially in non-first vowel slots.

The sequence **oï* can be found in the first vowel slot:
Word-initially: **oïra* 'near';
Non-initially in the first vowel slot: **hoï* 'forest', **koïna* 'behind', **noïr* 'sleep', **noïtan* 'wet';
Possibly **oï* occurs in non-first vowel slot in **sïroï* 'earth', but this may stem from earlier **sïruaï*.

The sequence **uï* can be found in the first vowel slot:
Word-initially: **uïla-* 'to cry';
Non-initially in the first vowel slot: **buïla* 'camel's nose plug', **huïtan ~ *hïutan* 'narrow', **kuï* 'sheath';
In non-first vowel slots: **alkuï* (?**alguï*) 'pot', **kanču*(*ï*) 'sleeve', **kïrguï* 'hawk', **muruï* 'crooked'.

The sequence **öi* may have been present in **göin* (?) 'shallow', **köisün* 'navel', **köiten* 'cold', **söike* 'earring'. The QG languages offer little evidence to distinguish **öi* from **üi*.

The sequence **üi* can be found in the first vowel slot:
Word-initially: **üile* 'work';
Non-initially in the first vowel slot: **büile* 'gums', **güi-* 'to run', **tüimer* 'steppe fire';
In non-first vowel slots: **tedüi* 'that much' (not in QG), **ügüi ~ *ügei* 'absent'.

3.13.3.2. Development

In Dagur **oï* and **uï* will merge into *oi*, as in Dag *nɔitun* < **noïtan* 'wet', *ɔis* < ?**uïsun* (other languages suggest **üisün*) 'birchbark'. However, **oï* and **uï* are subject to labial breaking when followed by **a*, as in *wair* < **oïra* 'near', *x*^{w}*ainə* < **koïna* 'behind'. Likewise **üi* has resulted in *ui*, but when followed by **e* it also underwent labial breaking, as in *wəil* < **üile* 'work'.[191]

In QG, the sequences **öi* and **üi* merged into **üi*. The complex vowels **oï, *uï, *üi* may monophthongise and merge to various degrees.

**oï* may result in *oi/ui* , or may be 'sinified' into > *uai* , or unrounded > *ai* > *ei* > *i:*. The following example illustrates the diversity of reflexes among and within languages: MgrH *fui:, fi:, fe:, xʊi,* BaoÑ *hi:,* Kgj *he,* Dgx *xoi* < **hoï* 'forest'. A

[191] Dag *wəil* could also be explained as a borrowing from Manchu.

deviating development in all of QG is that of **noïr* ‘sleep’, in which **oï* resulted in *ɔ:*. This represents an early regional irregularity.

In the QG languages **uï* and **üi* tend to become indistinguishable. Both **uï* and **üi* usually result in *ui* in Dongxiang and Mangghuer, while they are normally contracted into *ü:* in EYu, and into *i:* in Baoan and Kangjia.

3.13.4. Development of **au* and **eü*

3.13.4.1. Distribution

The sequence **au* can be found in all vowel slots:
Word-initially: **au* ‘wide’, **aula* ‘mountain’, **auškï* ‘lungs’;
Non-initially in the first vowel slot: **bau-* ‘to descend’, **daun* ‘sound’, **haul-* ‘to run’, **ǰau-* ‘to bite’, **ǰaura* ‘middle’, **kaučïn* ‘old’, **sau-* ‘to sit’;
In non-first vowel slots: **bïrau* ~ **burau* ‘calf’, **galǰau* ‘crazy’, **hasau-* ‘to ask’, **ǰalau* ‘young’, **kalaun* ‘hot’, **sïbaun* ‘bird’, **sïmaul* ‘mosquito’.

The sequence **eü* can be found in all vowel slots:
Word-initially: **eüden* ‘door’, **eülen* ‘cloud’, **eüs-* ‘to develop’;
Non-initially in the first vowel slot: **deü* ‘younger sibling’, **deüli-* ‘to jump’, **heür* ‘nest’, **ǰeüdün* ‘dream’, **ǰeün* ‘needle’, **seüder* ‘shade’, **seül* ‘tail’.
In non-first vowel slots: **bileü* ‘whetstone’, **ebčeün* ‘chest’, **ereün* ‘chin’, **hileü* ‘surplus’, **köbeün* ‘son’.

3.13.4.2. Development

Examples for the monophthongisation **au* and **eü* into **u:* and **ü:* are already documented in the Middle Mongol sources in Arabic script, but contraction was not a general rule. In central Mongolic as well as in Eastern Yugur this contraction was systematically executed.

In Dagur **au* and **eü* are typically preserved in the first vowel slot, as in *aul* < **aula* ‘mountain’, *əud* < **eüden* ‘door’. Elsewhere **au* has become *u:* or *o:*, without an obvious system, e.g. *xalu:n* < **kalaun* ‘hot’, *ǰalɔ:* < **ǰalau* ‘young’. Beyond the first vowel slot **eü* has resulted in Dagur *u:*, as in *ərču:* < **ebčeün* ‘chest’.

In Eastern Yugur, the normal developments of **au* and **eü*, regardless of position, are *u:* and *ü:*. The latter may be shortened to *ü* or ‘re-diphthongised’ into *ʉi*.[192]

Mongghul features diphtongs and contracted reflexes. This largely depends on the dialect, as shown in the following examples (here compared with Mangghuer).

[192] In Qinglong dialect, *ü:* may have been lowered to *ö:*, as in *dö:* < **deü* ‘younger brother’, *jö:dən* <**ǰeüdün* ‘dream’.

CM	Mongghul				Mangghuer	
	S/M	Khas	Donggou	Dongshan		
**bau-*	*bu:-*	*bau-*	*bɤu- ~ bau-*	*bɤu-*	*bau-*	to descend
**ǰau-*	*ʑiu:-*	*ʑau-*	*ʒɔu- ~ ʑau-*	*ʒɔu-*	*ʒau-*	to bite
**keü*	*ku:*	*ko:*	*ku: ~ kïu*	*kïu*	*kau*	son
**ebčeün*	*śʑiu:*	*śʑau:*	*śiʑɤu ~ śiʑou*	*śiʑou*	*śiʑɔu*	chest

As can be seen above, both **au* and **eü* may result in Mangghuer *au*. After some consonants **eü* resulted in *iau*, as in *diau* < **deü* 'younger sibling', *śauʑi* < **seüǰi* 'hip'. This can be seen as an attempt to make sense of the non-Chinese diphthong **eu* that must have existed in earlier Monguoric.

In the Baoanic languages only Dongxiang preserves the diphthongal pronunciations of **au* and **eü*. The former often becomes *ɔu* in the first vowel slot. As in Mangghuer, **eü* undergoes 'sinification' in Dongxiang, e.g. *ʑiau* < **deü* 'younger sibling', *śiauʑie* < **seüder* 'shade'. In Baoan and Kangjia, **au* and **eü* tend to merge into *u* or *u:*, but exceptions abound. Unlike Baoan, Kangjia retains diphthongs sporadically, as in *mu ~ mau* < **mau* 'bad', Kgj *sʉtər ~ səʉtər* < **seüder* 'shade', *devʉ* < **deü* 'younger sibling'.

Moghol preserved a diphthong (*öʉ* according to Ramstedt) in all positions, e.g. *öʉl-* < **haul-* 'to run'.

There are quite many words that deviate from the above. Aberrant developments shared by a number of dialects may be useful for taxonomic purposes. Shirongol idiosyncracies include the following:

CM	local form(s) suggest
**auškï* 'lungs'	Mongghul *o:* instead of *au* or *u:*
*?*eüken* 'fat'	Monguoric *o:* instead of *ɔu* or *u:*
**aula* 'mountain'	No diphthong anywhere in QG
**eülen* 'cloud'	No diphthong anywhere in QG
**eüden* 'door'	Kgj and Dgx suggest earlier **uiden* or **oiden*[193]
**kaučïn* 'old'	Baoanic suggests earlier **qoičin*
**(h)aur* 'anger; steam'	Baoanic suggests **hor*
**heür* 'nest'	All of Shirongol suggests **ho:r*
**seüder* 'shade'	Baoan suggests **seiter*
**seül* 'tail'	Baoanic suggests **sel*

In short, although the peripheral languages are more informative regarding these vowel sequences than central Mongolic, individual word shapes can be hard to evaluate.

[193] This is supported by Moghol, probably fortuitously.

3.13.5. Development of **ïa, *ie, *ïu, *iü*

3.13.5.1. Distribution

The diphthongs starting with **ï/*i* do not occur word-initially, unless words like **yasun* 'bone', **yeren* 'ninety' are analysed as **ïasun*, **ieren*. While this analysis is possible, it would not eliminate the need for a consonant **y* in the CM system, as intervocalic *-y-* would be needed in **bayan* 'rich', **beye* 'body'.[194]

The sequence **ïa* can be found:
Non-initially in the first vowel slot: **ǰïar(ï)* 'musk', **kïag* (not listed) 'couch grass', **nïa-* 'to glue'.
In non-first vowel slots: **ačïan* 'load', **arïa* 'molar, tusk', **dolïa-* 'to lick', **horïa-* 'to wind', **kurïa-* 'to gather', **marïan* 'fat', **sunïa-* 'to stretch oneself', **takïa* 'chicken'.

The sequence **ie* can be found:
Non-initially in the first vowel slot: **sie-* 'to pee'.
In non-first vowel slots: **belčier* 'pasture', **bülien* 'lukewarm', **ebsie-* 'to yawn', **hinie-* 'to laugh', **hülie-* 'to blow', **kerie* 'crow', **tülien* 'firewood',**ünien* 'cow'.

The sequence **ïu* can be found:
Non-initially in the first vowel slot: **hïutan* ~ **huïtan* 'narrow', **kïursun* 'nit', **nïu-* 'to hide', **nïun* 'boy', **nïur* 'face'.
In non-first vowel slots: **arčïur* 'towel', **arïun* 'clean', **bïlǰïur* ~**bïlduur* (?PM **bïldïur*) 'small bird', **bïsïun*, **čakïur* 'lighter', **čïul-* 'to gather', **gasïun* 'bitter', **horčïul-* 'to turn (tr.)', **kosïun* 'beak'.

The sequence **iü* can be found:
Non-initially in the first vowel slot: **niül* 'sin' (not listed), **siü-* 'to strain', *?*siüči* ~ *?*seüči* 'chisel', **siüderi* 'dew', **siür-* 'to sweep', **siüsün* 'juice'.
In non-first vowel slots: **delbiür* 'fan', **deliün* 'spleen', **ebčiün* 'chest', **eriün* 'chin', **gesiün* 'branch', **seriün* 'cool', **teriün* 'head'.

3.13.5.2. Development

The most common change affecting the diphthongs starting with **ï/*i* is the assimilation of this first element to the second. This type of development can already be observed in Middle Mongol sources.[195] In general Buriat, Khamnigan and Dagur preserve most instances of **ï/*i* in the form of palatalisation of the preceding consonant. Retention of such palatal elements is also found in Khalkha, Ordos and

[194] The analysis of these words as **baïan* 'rich', **beie* 'body' is also undesirable, because it would introduce a new type of vowel sequences.

[195] The assimilation of complex vowels was already possible prehistorically, giving rise to modern regional variants such as **bïlǰïur* ~ **bïlduur* 'small songbird' and **kaǰïar* ~ **kadaar* 'bit of a bridle'.

Kalmuck. Several Shirongol languages preserve some of these diphthongs as well, but they were completely lost in Ñantoq Baoan. The diphthongs must have been present both in Proto Baoanic and Proto Monguoric. Eastern Yugur has consistently monophthongised both **au* and **eü*.

Assimilation of the diphthong elements may have taken place in any period. In some cases it must have been quite early because it affected the development of the consonants, for instance Dag *səːs* 'urine' apparently developed from an assimilated form **seesün* rather than from the CM **siesün* suggested by the majority of Mongolic languages. The assimilation has prevented the development **s- > š-* that would have taken place preceding **i*. Other such cases include *gasuːn* < **gasïun* 'bitter' and *gəsuː* < **gesiün* 'branch' (see 4.6.1.).

Likewise Khalkha *zaar* 'musk' stems from the assimilated form **ǰaar* rather than from the earlier **ǰïar*. The original diphthong would have resulted in a Khalkha form with *ǰ-*.

Examples for the development of the complex vowels **ïa* and **ïu*:

CM	Bur	Dag	MgrM	EYu	
**nïa-*	*nyaa-*	*nʸaː-*	*nia-*	*naː-*	to glue
**kanïa-*	*xanyaa-*	*xanʸeːd-*	*qania-*	*χanaː-*	to cough
**kubïa-*	*xubaa-*	*xɔbʸeː-*	*quba-*	*χwaː-*	to divide
**karïa-*	*xaraa-*	*xaraː-*	---	*χaraː-*	to scold
**ǰarïa*	*zaryaa*	*ǰerʸeː*	---	*ǰaraː*	hedgehog
**horïa-*	*oryoo-*	---	*xoro-*	*χɔrɔː-*	to wind
**korïan*	*xoryoo*	*xɔrʸeː*	*qoraŋ*	*χɔrɔːn*	enclosure
**nïu-*	*nyuu-*	*nɔː-*	*niu-*	*nuː-*	to hide
**arïun*	*aryuun*	*aruːn*	*ariun*	*aruːn*	clean

Moghol also preserves a number of the diphthongs starting with **ï/*i*, as in *nioː-* < **nïa-* 'to glue'. Others are preserved indirectly, in that the element **ï/*i* has moved to the preceding vowel slot, e.g. Mog *qaïnʌ-* < **kanïa-* 'to cough', *ǰäira* < **ǰarïa* 'hedgehog', *ɯinà* < **ünien* 'cow'.[196]

In Dahejia Baoan, which does not generally preserve vowel length, the contraction lengths can be shortened and reduced to *ə*, as in:

BaoD	BaoÑ	CM	
arə-duŋ	*ariya*[197]	**arïa*	molar
χorə-	*hərə-*	**horïa-*	to wind

[196] *innaː-* < **hinie-* 'to laugh' may go back to a metathesised form *iïnaː-*, which would explain the unusual gemination of *-n-*.

[197] The Dahejia Baoan form is a compound with *duŋ* < **sidün* 'tooth'. The Ñantoq Baoan form need not be a relic of the original CM **arïa*, as it could also represent Baoan *arə* + Chinese *yá* 'tooth'.

The long vowel has been lost altogether in BaoX *χandoŋ* < **kanïa.dun* ‘cough’.

**kïursun* ‘nit’ shows irregular developments in Dag *k^{y}aur* (suggesting **kïura*), EYu *χairsən* (suggesting **kaïrsun*), MgrH *ćirʒə* (suggesting **kïrsun*).

Buriat, Dagur, and Khamnigan frequently preserve palatal elements in words with original **ie* and **iü*, but assimilations can be found even there, as in Dag *xinəːd* < **hinie.dün* ‘laughter’, Bur *ebheel-* < *?*ebsie.le-* ‘to yawn’; **keǰie* ‘when’ developed as if from **keǰee* in all three literary languages.

In the QG languages the diphthongs **ie* and **iü* have typically become indistinguishable from **ee* and **üü*.

3.13.6. Metathesis of diphthong elements

Some irregular forms, and resulting irregular correspondences between modern languages are due to the metathesis of diphthong elements. Several such cases of metathesis are actually documented in Middle Mongol, e.g. **hïutan* ~ **huïtan* ‘narrow’, both of them surviving into modern Mongolic, cf. Dag *x^{y}aut* from the former, and MgrM *xuitaŋ* from the latter variant.

**hïruar* ~ **hïraur* ‘bottom’ co-occur in Middle Mongol, but modern forms generally stem from **hïroar*, a development of the first variant. **ǰïloa* (< **ǰïlua*) ‘reins’ appears in sources in Arabic script as **ǰïlau*, **ǰïroa* ‘ambler’, as reconstructed on the basis of modern languages, appears in Arabic script as **ǰorïa*.[198] **komaul* ‘horse dung’ is documented early, but the variant **komual* has given rise to the modern forms in the central languages. Likewise in front-vocalic words there must have been alternations **üe* ~ **eü*, the former developing into **öe*. Dag *ǰəulə́ːn* ‘soft’ and *əuɣ^{w}* ‘grease’ suggest CM forms **ǰeülen* and **eükün*, whereas central Mongolic as well as the Shirongol languages suggest **ǰöelen* < **ǰüelen* and **öekün* < **üekün* (cf. MgrH *ʒ́oːlon, oːko*).

Dag *suidur* must be from **süideri* instead of **siüderi* ‘dew’; the metathesised form has prevented the development **s-* > *š-*.

The transfer of the diphthong element **ï* or **i* to another syllable is best known from Moghol (see 3.13.5.), but incidental cases can be found in central Mongolic:

Khalkha *malia* suggests **malïa* < **mïlaa* < **mïnaa* ‘whip’.
Khalkha *myaraax* reflecting **mïraa-* instead of **marïa-* ‘to spy on’
Ordos *mira:* reflecting **mïraan* instead of **marïan* ‘flesh’

3.14. Non-contraction vowel-lengths

The following sections will discuss some of the types of vowel lengths in Dagur, Eastern Yugur, and the Shirongol languages, which are not the result of contractions. The majority of these lengths are due to accent-related lengthening or other kinds of by-products of word structure.

[198] This is actually the form that was expected on the basis of the Turkic cognate **yorïga*.

3.14.1. Ill-spelled contraction lengths?

As mentioned above, publications on Mongolic long vowels have often distinguished between 'secondary' and 'primary' vowel lengths. In practice, Written Mongol spelling was used to determine in which category a vowel length belonged. A long vowel in the modern languages, which is written in Written Mongol by means of a *VCV* sequence (nearly always *VɣV* or *VgV*), was supposed to be a 'secondary' length resulting from contraction. A modern long vowel which is written with a single Written Mongol vowel letter was considered to be a 'primary' length, predating the contraction lengths.

Unfortunately, Written Mongol spelling has never been completely stable, and many words appear in two or more spelling variants. Through the centuries Written Mongol spellings were often influenced by the diverging spoken languages they represented. It goes without saying that not all recorded spellings are etymologically correct. Spelling variation is often informative, in that some of the variants are clearly attempts to represent changing pronunciations in the spoken idioms of the users of the script.

Ramstedt (1902) already stressed that *VɣV* and similar sequences in Written Mongol were not necessarily historically correct, and that these spellings were also used to write long vowels in loanwords.[199] In spite of this, many later scholars have used Written Mongol spelling for reconstruction purposes with insufficient caution.

A number of words with a long vowel in several modern languages are often or even exclusively spelled with a single vowel sign in WM, e.g. **haaga* spelled *aɣa* 'bran', **öer* (?) spelled *ör* 'dawn'. This seems to indicate that these vowels are either 'primary' or due to a type of lengthening different from contraction. However, such words may actually be contraction lengths which acquired a historically incorrect orthographical variant as their standard Written Mongol spelling. The following words, with long vowels supported by several languages, seem to belong to this group of miss-spelled words (historically correct spelling variants that are documented for several of these words are omitted here):

WM	CM as based on modern languages	
aɣa	**haaga*	bran, chaff
bolod	**boluad*	and
ör	**öer*	dawn
čaɣan	**čagaan*	white
čüče	**čüüče* < **siüče*	chisel
gegen	**gegeen*	bright
saral	**saaral*	grey (of animal's coat)
sina	**sïnaa*	wedge
sü(*n*)	**süü* ~ **sün* ~ **üsün*	milk
ula	**huula*	tinder
ü-	**hüü-*	to rot/stink

[199] Two examples for this practice that bear repeating are *byarwaas* written *birabaɣasun* < Ru. *perevóz* 'transfer' and *yaarmag* written *yaɣarmaɣ* < Ru. *yármarka* 'fair'.

The last word may also be a case of lengthening of the final vowel of a monosyllable (see 3.14.3).

The lengths are not systematically supported by all languages but are found in Khalkha, Buriat, Dagur, and incidentally in the QG languages:

Kh	Bur	Dag	EYu	MgrH	CM
aaga	*aaga.han*	*xa:ɣə*	---	---	**haaga*
süü(n)	*hün*	*su:*	*hsʉn*	*sun*	**süü*
uul	*u(u)la*	*x^{w}a:l*	---	*fula*	**huula*
ü-	---	*xu:-*	*hü:-*	*fu:-*	**hüü-*
üür	*üür*	*ur*	*ö:r*	*o:r*	**öer*

Cf. Kalm *üsn* ‘milk’, *ör* ‘dawn’, *ul* ‘tinder’. EYu *hsʉn* ‘milk’ and possibly MgrH *sun* may derive from the variant **üsün* (as does the Kalmuck form) rather than from a form **sün* or **sü:*.

In case of variant spellings[200], the modern languages usually make it possible to determine which spelling is etymologically correct (which may be confirmed by Turkic cognates).

WM	CM	CT	
daɣa(*n*) ~ *daɣaɣa*(*n*)	**daagan*	**yapak*/**yapagu*	colt
čoqor ~ *čooqor*	**čoakar*	**čupar* (?=)	variegated[201]
kö ~ *köge* ~ *kögege*	**köe*	**kög*	soot in a pan

In the case of Written Mongol *daɣa*(*n*)/*daɣaɣa*(*n*) ‘colt’, the modern forms such as Kh *daaga*(*n*), EYu *da:ğan*, MgrH *da:xa*, Dag *da:ɣ* point to a CM form **daagan*. Thus the short spelling *daɣan* reflects neither etymological reality nor the pronunciation, while *daɣaɣan* is correct within the Written Mongol spelling rules, with the first *ɣ* now denoting vowel length, and the second still pronounced.

In the case of *kö* ~ *köge* ~ *kögege*, *kö* is based on the modern Khalkha pronunciation *xo:*, though it fails to represent the long vowel. *Köge* is an etymologically correct spelling. *Kögege* is a spelling apparently meant to indicate the long vowel, but contains one *-g-* too many from an etymological viewpoint. Likewise Written Mongol *daɣaɣa-* for **daa-* ‘to bear’ and Written Mongol *egege-* for **hee-* ‘to warm in the sun’ contain a superfluous syllable.

---

[200] Note that Written Inner Mongolian spelling seems to have adopted as standard the shorter spelling variants in many such cases, e.g. *čaɣan* ‘white’, *gegen* ‘bright’, *daɣa* ‘foal’, *kö* ‘soot’, *čoqor* ‘speckled’, etc. Cf. also the spelling *buqur* for Kh *buxuur*, Dag *bɔɣɔ:r* ‘buttocks’.

[201] The traditional equation of CM **čoakar* with CT **čupar* is morphologically problematic. Since most Mongolic words lack a Turkic cognate, it is often difficult or impossible to tell the etymologically correct spellings from the incorrect ones, and to determine the nature of a Mongolic vowel length. (Ordos supports **čoakar*; EYu favours **čoakur*)

The vowel length of Khalkha *buurcag* 'beans' is only attested in central Mongolic; that of Khalkha *büüreg* 'flea' is supported by Kalmuck and one of the variants attested for Mongghul.[202]

Some lenghts that (based on Written Mongol spelling) are considered to be contraction lengths may be primary lengths, at least it may sometimes be impossible to confirm that a written *g*, *ɣ*, etc, represents an actual etymological consonant.

In a small number of cases the corresponding Turkic form actually suggests that there never was a consonant, e.g. Khalkha *uuc* 'back', Written Mongol *uɣuča*, which seems to stem from Turkic **u:ča*. There are some other instances where the Written Mongol intervocalic consonant probably does not reflect a real consonant: Khalkha *toodog*, Written Mongol *toɣodoɣ* 'bustard', if corresponding to Turkic **to:d* 'id', Khalkha *xöörög*, Written Mongol *kögerge* 'bellows' if corresponding to Turkic **kö:r(ü)k* 'id', Khalkha *čiig*, Written Mongol *čigig* 'moisture' if corresponding to Turkic **čï:g* 'id'. If these comparisons are correct, the Written Mongol spellings can only be interpreted as spelling conventions.

As in Old Uigur (the source of many conventions of Written Mongol spelling), we find words in which the grapheme <*u*> is doubled, although the spoken languages lack evidence for a long vowel. Perhaps it was a spelling convention used to indicate *o* rather than *u*, but this explanation does not always make sense. Perhaps the doubling is just a graphical means to avoid confusion with other words, much like the accents in Romance languages, e.g. French *ou* ≠ *où*, *a* ≠ *à,* etc: Examples include Written Mongol *ɣool* 'valley', *qoora* 'poison', and *toor* 'net'. One case where the spelling *oo* corresponds to an actual long vowel in QG is Written Mongol *doora* 'underside' (which was shortened in Dagur and central Mongolic).

For the (lack of) correlation between Mongolic and Turkic vowel length see further 3.15. below.

3.14.2. Non-contraction vowel lengths in peripheral Mongolic

In the following pages those vowel lengths will be discussed that are apparently not 'secondary', i.e. the result of contraction of two consecutive short vowels after the elision of an intervocalic consonant. As will become clear, part of the non-contraction lengths could have arisen relatively recently in certain phonetic environments, while others remain unexplained. Examples from the latter group may reflect actual CM vowel length (as the label 'primary' vowel length suggests), or be explained as secondary developments by future research.

Interestingly, Poppe concluded that there were original long vowels in Mongolic, and Doerfer concluded that their existence can not be proved, both on the basis of largely the same evidence. In the following paragraphs, some of the well-known evidence will be presented in a different way, and combined with new evidence from the peripheral languages. This includes a discussion of materials of Eastern Yugur, which was hardly described at the time when Poppe, Doerfer, Ligeti

[202] In these two words the CT cognates **burčak* and **bürge* suggest that these are short vowels which for unclear reasons have been secondarily lengthened in Mongolic, rather than contraction lengths.

and others discussed Mongolic vowel length. The focus will be on the various ways of vowel lengthening which may have occurred.

The peripheral languages, especially Dagur and Mongghul, show numerous vowel lengths that have not been observed elsewhere. In a number of cases both Dagur and Mongghul display length in the same word. Some correspondences make it tempting to reconstruct the length for CM. Perhaps the best-known and most striking examples for 'primary' long vowels in Mongolic are the following two:

WM	Dag	EYu	MgrH	CM	
modu(n)	*mɔ:d*	*mu:dən*	*mo:də*	**modun* (?**mo:dun*)	wood; tree
tabu(n)	*ta:wu*	*ta:βən*	*ta:vun*	**tabun* (?**ta:bun*)	five

There are some other words in which Mongghul and Dagur agree, but for which the Eastern Yugur materials are ambiguous:

WM	Dag	EYu	MgrH	CM	
temeče-	*təmče:-*	*temče-*	*tənʒ́e:-*	**temeče-*	to struggle
uya-	*xuya:-*	*ɣ́a:-, hia-*	*fuya:-*	**huya-*	to tie up
umba-	*xumpa:-*	*mba-*	*xumba:-*	**humba-*	to swim
örögel	*ərɣ^{w}lə:*	*örlö(:)*	---	**öreele*	hobble

There are occasional supporting forms in other languages, such as BaoÑ *mba:-* 'to swim', and Buriat *uyaa-* 'to tie up'.

Despite the lack of systematic support from all three languages, the long vowel in these words is still found in two peripheries, and at first sight make a quite convincing case for CM vowel length. That is, they seem to suggest CM reconstructions **temeče:-*, **huya:-*, **humba:-* rather than the forms listed above. However, the striking Dagur-Mongghul parallels are outnumbered by far by instances where length is only found in one of the two languages, and cannot be included in the CM reconstruction form. Examples for disagreement between Dagur and Mongghul:

Dagur	Mongghul	CM	
bərɣə:n	*urgen*	**berigen*	sister-in-law
šil	*śilo:*	**silön*	soup
yaləɣ-	*ləġa:-*	**ïlga-*	to choose
wail-	*ula:-*	**uïla-*	to cry
kuitur-	*kuidəre:-*	**kuitere-*	to become cold
gal^{y}da:-	*ġalda-*	**galda-*	to make a fire
xaičla:-	*xaiʒ́ila-*	**kaïčïla-*	to clip
šiŋgə-	*śiŋge:-*	**siŋge-*	to be digested
šiŋgə:n	*śæŋgan*	**siŋgen*	thin

Again, comparison with the Eastern Yugur forms, does not lead to a regular picture, e.g. EYu *be:rge* ‘sister-in-law’ has a vowel length in the first syllable instead of the second, which may be due to the influence of *beːrə* ‘daughter-in-law’ (**beri*). The many other words with contradicting forms in Dagur, Eastern Yugur, and Mongghul, include:

WM	Dag	EYu	MgrH	CM	
yara	*yar*	*ya:ra*	*ya:ra*	**yara*	wound
kekere-	*kəkreː-*	*ge:gəre-*	*kəgəreː-*	**kekere-*	to belch
tulɣa	*t*^{w}*a:ləɣ*	*tu(:)lġa*	*tulġa*	**tulga*	support

3.14.3. Lengthening of the final vowel of monosyllables

In Dagur the final vowel of CM monosyllabic stems is lengthened. There are examples of stems in **a*, **i*/**ï*, **u*, **ü*. There are unsystematic equivalents in central Mongolic. Examples:

CM	Dagur	length also found in	
**a-*	*a:-*	MgrH, central **a:-dal*	to dwell, to be
**ba*	*ba:*		we (excl.)
**ta*	*ta:*	Brg *ta:*	you (pl.)
**bï*	*bi:*	Brg *bi:*	I
**čï*	*ši:*	Brg *ši:*	you (sing.)
**ki-*	*xi:-*	Khalkha *xiy-*, EYu *ki:-*	to do; to pour
**bü*	*bu:*	Khalkha *büü*, MgrH *bi:*	don’t
nu-* (?no-*, ? **onu-*)	*nɔ:-*	EYu *nu:-*	to hit

Dagur *a:n* < **a* ‘they’and *i:n* < **ï* ‘s/he’ may also belong here, as these pronouns owe their final *-n* to the oblique forms. The vowel length may date from before this addition, or may be a later development under the influence of the other personal pronouns.[203]

The vowel length in Dag *su:* (?< **sü*) ‘milk’ also has equivalents in some other languages, this word elsewhere has the shape **sün* or **üsün*.

EYu *χa:* < **ka* ‘upper front leg’ also corresponds to vowel lengths in central Mongolic, but this word is absent from Dagur.

In addition to the prohibitive particle **bü*, Monghhul has also lengthened the final vowel of the negation particles **ese* and **ülü*, resulting in *si:* and *li:*, both with unexpected change of the vowel quality to *i*.

[203] Interestingly the oblique stems of *ba:* ‘we (excl)’ and *ta:* ‘you (pl.)’ are *ma:n-* and *ta:n-* with long vowel, while *bi:* ‘I’ and *ši:* ‘you (sing.)’ have oblique stems *min-* and *šin-* with short vowel. Of course the vowels here lack the word-final position that supposedly caused the vowel to be lengthened.

Apparently monosyllabic stems in **e*, **o*, or **ö* were rare or non-existent. But **nu-* ~ **onu-* may have been**nɔ-* (neither of these proposed CM forms would solely be able to explain all modern forms).

3.14.4. Accent-related lengthening in the final syllable

In Mongghul the accented last vowel of word stems is often lengthened. The same happens, less frequently, in Eastern Yugur and Ñantoq Baoan. This phenomenon is limited to words of a certain structure. It occurs predominantly in the final syllable of disyllabic stems ending in a non-high vowel or in disyllabic stems ending in **n* or **l* preceded by a non-high vowel. The first syllable may have a high or non-high vowel. This word structure, which triggers lengthening of the second vowel in Mongghul, is the same structure that causes the accent to fall on the second vowel in Eastern Yugur. In this regard this phenomenon is largely complementary with that described for Eastern Yugur in 3.14.5., involving the lengthening of the non-high vowel of the *first* syllable. However, not all words with this structure undergo lengthening.

Only some vowel lengths in this position in Eastern Yugur or Ñantoq Baoan are not confirmed by Mongghul, such as EYu *ʉde:* < **üde* 'noon', EYu *ywe:* < **üye* 'generation'. Agreement between all languages is rare. In some cases two languages may agree:

MgrH	BaoÑ	CM	
nta:-	*ta:-*	**unta-*	to sleep
ula:-	*la:-*	**uïla-*	to cry
una:-	*na:-*	**una-*	to descend
uda:-	*nda:-*	**uda-*	to be late
daġa:-	*daġa:-*	**daga-*	to follow
xumba:-	*mba:-*	**humba-*	to swim

Dagur occasionally agrees with Mongghul, as in **huya-* 'to tie up' (see above), but such similarities are greatly outnumbered by forms that disagree between the two languages, and may be coincidental.

Mongghul stands alone in most cases (short variants may exist but are not listed here). However, this development is still not regular enough to call it a sound law. Even in Mongghul there are many words of a similar structure in which the last vowel was not lengthened, e.g. *ula* 'sole of the foot' (**ula*), *ule* 'work' (**üile*), *ugo* 'word' (**üge*).

Here follows a, far from exhaustive, list of Mongghul examples.

Nouns ending in a non-high vowel:

MgrH	EYu	CM	
dusa:	*tusa*	**tusa*	profit
fure:	*hʉre*	**hüre*	seed

(continued)

MgrH	EYu	CM	
ġada:	*ġada*	**kada*	rock
ġuda:	*ġuda*	**kuda*	male in-law
kəde:	*kede*	**kete*	steel for flint
mula:	*məla*	[regionalism]	small
noġdo:	*nɔgtɔ*	**nogta*	halter
turma:	*turma*	**turma*	turnip
xura:	*xura*	**kura*	rain
ʒarma:	---	**ǰarma*	hail
ʒila:	*ǰula*	**ǰula*	lamp

Nouns ending in *-l or -n* preceded by a non-high vowel:

gəre:l	*gerel*	**gerel*	light (noun)
ra:l	*aral*	**aral*	shaft
ra:l	*aral*	**aral*	river valley

MgrH	EYu	CM	
baya:n	*bəyan*	**bayan*	rich
dura:n	*dura*	**dura*	liking
kəʒe:n	*kəčə*	**keče*	slope
kurge:n	*kʉrgen*	**küregen*	brother-in-law
lumo:n	*manan* (?=)	**noman*	mole (animal)
nama:n	*manaŋ*	**manan*	mist
nʒe:n	*eǰen*	**eǰen*	self
noyo:n	*niyɔn*	**noyan*	lord
rme:n	*örmön*	**öreme*	skin on milk
səme:n	---	**süme*	temple
ʂdogo:n	*hʉtgwen*	**hödken*	dense
śilo:	*šelen ~ šölön*	**silön*	soup
śime:n	*šəme*	**sime(n)*	juice
təme:n	*temen*	**tümen*	ten thousand

Some words with a similar structure in Mongghul may involve compensation length, e.g. MgrH *ġura:n* 'three' from CM **gurban* after the loss of the *-v-* < **b*. In spite of the similar structure, a different development is found in MgrH *de:ren* 'four' < CM **dörben* where the first vowel was lengthened, and *xaran* 'ten' < CM **harban* where the **b* was lost without lengthening either of the vowels.

Incidentally, the preservation, and in some cases the addition, of final *-n* in the nouns is unexpected, as the **-n* of nouns is nearly always lost in Mongghul.

Verbs ending in a non-high vowel:

MgrH	BaoÑ	EYu	CM	
bula:-	*bəla-*	*bula:-*	**bula-*	to bury
bura:-	*wara-*	*bara-*	**bara-*	to finish
ćina:-	*ćina-*	*čəna-*	**čïna-*	to cook
daġa:-	*daġa:-*	*dağa-*	**daga-*	to follow
dava:-	---	*daβa-*	**daba-*	to cross
fule:-	---	*hele-*	**hüle-*	to remain
fuya:- ~ śa:-	*χa:-* (?=)	*xa:-, hia-*	**huya-*	to tie up
nəme:-	---	*neme-*	**neme-*	to add
ntəra:-	---	*ndara-*	**untura-*	to sleep
nura:-	---	---	**nura-*	to collapse
sune:-	---	---	**söne-*	to go out (fire)
śʒe:-	*ʂʒe-*	*hče-*	**hičə-*	to be ashamed
śiŋge:-	---	*šeŋge-*	**siŋge-*	to digest
śiġa:-	*xiχa-*	*šəqa-*	**sïka-*	to squeeze
śira:-	*xira-*	*šəra-*	**sïra-*	to roast
təbde:-	---	*debte-*	**debte-*	to soak (intr)
ula:-	*la:-*	*ü:la-*	**uïla-*	to cry
una:-	*na:-*	*na:-*	**una-*	to descend
xambura:-	*hamara-*	*aməra-*	**hamura-*	to rest
yo:-	---	---	**oya-*	to sew
ʒira:-	*ʒira-*	*ǰü:ra-*	**ǰuura-*	to mix[204]

Among the Mongghul verbs with lengthened final vowels we also find the original causatives *kurge:-* < **kürge-* 'to bring, send, deliver' from **kür-* 'to reach, arrive', and *surġua:-* < **surga-* 'to teach' from **sur-* 'to learn'. These verbs are perhaps no longer considered as causatives; in fact there are new Mongghul causative forms based on **kür-* and **sur-*, with the meanings 'to cause to reach' and 'to cause to learn' (MgrH *kurġa-* and *suruġa-*). The vowel of the productive causative suffixes *-ġa, -lġa* is not long.

Incidentally, verbs with lengthened final vowels did not merge with their own causatives, as these have been provided with newly added causative endings. *xada:-* 'to dry (intr)' represents CM **kata-*, and the function of the causative **kataa-* is now fulfilled by a new formation *xada:-lġa-* 'to dry (tr)'.

Secondary lengthening may contribute to the convergence of originally different shapes. In Junast's data MgrH *təme:n* 'camel; ten thousand' represents both **temeen* 'camel' and **tümen* 'ten thousand'.

There are several discrepancies between Eastern Yugur sources with regard to length of non-high vowels. The forms documented by Junast more often have long final vowels reminiscent of Mongghul, including the following.

[204] This example could be seen as a case of length metathesis, but it may simply be an accidental co-occurrence of shortening of a double vowel and lengthening of the final vowel.

Junast	Bolčuluu	CM	
soġdoː-	*sɔgtɔ-*	**sogta-*	to get drunk
tele:-	*tele-*	**tüle-*	to burn
törö:-	*törö-*	**töre-*	to be born
χaġaraː-	*hqara-*	**kagara-*	to break (intr.)
mere:n	*mʉren*	**mören*	river
pəseː	*pəsəi*	**büse* (?**büsei*)	belt

Lengthening of original high vowels is uncommon. Lengthening of CM **i* is found in Mongghul in *puśiː* 'wrong' < **bisi* 'not', *pugəliː* < **büküli* 'all', *daliː* (~ *dalə*) < **adali* 'similar'.[205]

The negative particles MgrH *biː* (**bu/*bü*) 'don't', *liː* 'not' (**ülü*), *siː* 'not' (**ese*) all replaced their original vowels by an irregular *iː*, probably due to group analogy.

There seem to be hardly any common nouns and verbs with this development. MgrH *daːliː* 'shoulderblade' may go back to a form ending in a complex vowel **daluï* instead of the commonly reconstructed **dalu*; a form **daluï* could account for some of the other Shirongol forms as well.

[205] Some of these function words may owe their long vowel to the fact that they have incorporated the copula *iː*.

3.14.5. Accent-related lengthening in non-final syllables

In Eastern Yugur and Mongghul one finds numerous long vowels in the first syllable which are short in the central languages and Dagur. Those Baoan dialects that preserve phonemic length, notably Ñantoq Baoan, also feature lengthening in some words of this structure. These lengths are unexpected since the accent in the QG languages generally falls on the final syllable, which as seen in the previous paragraph is another source of vowel lengthening in Mongghul and Eastern Yugur.[206]

However, Poppe (1955:26 and 73-75) recognised a certain pattern in the Mongghul words with vowel lengthening in the first syllable. He noticed that many of the 'affected' stems have the structure *(C)VCV(C)*, in which the first vowel is non-high and the second high. A similar pattern applies to many Eastern Yugur words with lengthened vowels. Virtually all cases are disyllabic stems with a non-high vowel in the first syllable and a high vowel in the second. In Mongghul the first syllable of the words in question is open, while in Eastern Yugur it may either be open or closed. The similar structures of the relevant words suggests that we are dealing with a conditioned development rather than the preservation of ancient vowel lengths. As high vowels in general tend to be slightly shorter than non-high vowels, it is likely that the vowel lengths in these words are due to a kind of compensatory lengthening.

Bolčuluu & Jalsan (1990: 142-143)[207] describe how the vocalism has given rise to deviations from the final accent rule in Eastern Yugur. Synchronically the accent tends to avoid the vowel *ə* (which in Eastern Yugur mostly replaces the four CM short high vowels in non-first syllables), if a non-high vowel is also present in the stem.

The default accent placement is on the last syllable. This is found in stems containing only short non-high vowels such as *al'tan* (**altan*) 'gold', *hele'ɣen* (**heligen*) 'liver', and in stems containing only short high vowels, such as *pə'čəg* (**bičig*) 'writing', *xəmə'sən* (**kïmusun*) 'nail'. However, if the stem contains a contraction length or diphthong, that will take the accent (vowel height is irrelevant), e.g. *ǰɔ'rɔ:* (**ǰïroa*) 'ambler', *tɔlɔ'ğɔi* (**tolagaï*) 'head'; *dʉ're:* (**döree*) 'stirrup', *du'la:n* (**dulaan*) 'warm', *xa'ru:l-* (**karïul-*) 'to reply', but *'ki:re* (**keere*) 'outdoors', *'ǰe:ren* (*ǰeeren*) 'Mongolian gazelle'. If the stem contains long vowels only, the accent will fall on the final one.[208] If the stem contains short high and non-high vowels, the final accent will be on the last syllable if that has a non-high vowel. If *ə* is in the final syllable of a disyllabic stem, the accent will avoid the *ə*, and fall on the vowel of the first syllable in, as in *'ğatən* (**katun*) 'empress', *'xɔ:nə (*konïn)*

---

[206] Related to this may be the lengthening of non-high vowels before the suffix **-sUn*, a phenomenon found in both central and peripheral languages.

[207] In Malov's limited Eastern Yugur materials, edited by Kotwicz, non-final accent seems to be indicated with the grave accent in a handful of words. Junast (1981:14) only mentions the basic rule that accent is word-final.

[208] Note Bolčuluu & Jalsan (1990:143) *'xɔ:nə* 'sheep'> accusative/genitive *xɔ:'ni:n*, reflexive possessive *xɔ:nə'ya:n* 'one's own sheep', where the long suffix vowel takes over the accent. Compare the dative-locative *'xɔ:nədə*, where the accent remains on the stem.

‘sheep’. In trisyllabic and longer stems, the accent will fall on the rightmost non-high vowel available, as in *he'nesən* (**hünesün*) ‘ash’, *'belwəsən* (**belbisün*) ‘widow’, *jə'ğasən* (**jagasun*) ‘fish’, *xulă'ğaičə* (**kulagaïčï*) ‘thief’.

Now in many words this accent shift has led to the lengthening of the vowel that came to bear the accent, although this is not obligatory, see *'ğatən* ‘empress’ vs. *'xɔ:nə* above. Many words have been recorded both with short and long vowel, e.g. EYu according to Bolčuluu & Jalsan *'sdasən* ‘thread’ vs. Junast’s *hda:sən.* (**utasun*). Other examples of which Bolčuluu & Jalsan explicitly mark the accent are *'mu:dən* (**modun*) ‘wood’, *'dɔ:lə-* (**dolïa-* via intermediate QG form **dolï-*) ‘to lick’. Although accent is not indicated in the Eastern Yugur vocabularies, the same accent pattern should equally apply to the other examples listed below.

For Mongghul, all authors only note word-final accent. The example *na:'də-* < **naad-* ‘to play’, stressed on the *ə* (which in this case is not even original, but a connective vowel reinterpreted as part of the stem), contrasts with the accent rules of Eastern Yugur. In Ñantoq Baoan final accent is also the norm, e.g. *mə'rə* < **morïn* ‘horse’. The fact that Mongghul and Eastern Yugur have developed this type of vowel length under the same circumstances, albeit usually not in the same stems, suggests that, at an earlier stage of its development, the Mongghul accent may have behaved like that of Eastern Yugur.

Unfortunately the evidence is not at all consistent in the languages concerned. In both Mongghul and Eastern Yugur there are many words with a similar phonetic structure which did not develop a long vowel. Furthermore, in words which are preserved in both Mongghul and Eastern Yugur, the forms rarely agree. In fact the disagreements seem suspiciously systematic, which may indicate that another factor, as yet undiscovered, has played a role in shaping these forms. E.g. **konïn* ‘sheep’, and **morïn* ‘horse’ are long in Eastern Yugur, but short elsewhere in Mongolic; **hödün* ‘feather’ is only long in Mongghul; **hodun* ‘star’ is long in Mongghul and Ñantoq Baoan, but short elsewhere.[209] Such contradictions make it impossible to reconstruct length in these words for a common ancestral language.

CM **modun* ‘tree’ and **tabun* ‘five’, the most widely accepted examples of primary vowel length in CM, are the only cases supported by three peripheral languages. They may be attributed to secondary vowel lengthening.

CM	E. Yugur	Mongghul	Dagur	
**modun*	*mu:dən*	*mo:də*	*mɔ:d*	wood, tree
**tabun*	*ta:βən*	*ta:vun*	*ta:wu*	five

The first example is complicated by the fact that QG languages may be influenced by Ch *mùtou* ‘wood’ and Dagur by the Tungusic word *mɔ:* ‘tree; wood’. The length in the second example is more difficult to explain. The length could have arisen to maximise the contrast with **tabïn* ‘fifty’. It is not clear whether this could have been

[209] Malov (Kotwicz) indicates accent on the first syllable in **hodun* ‘star’ for Eastern Yugur.

an early development, later neutralised again in central Mongolic, or an independent development in Dagur and the QG languages.

Below follows a list of the Mongghul and Eastern Yugur words of the structure (C)VCV in which lengthening occurred. Long vowels may appear in either Eastern Yugur or in Mongghul, or in both.[210] Apart from **modun* 'tree' and **tabun* 'five', Dagur does not confirm these lengths.

CM	E. Yugur	Mongghul	Dagur	
**alï*	*a:lə*	*alə*	*alʸ*	which
**amun*	*amən*	*a:mu*	*am*	millet[211]
**daku*	---	*da:xu*	*daw*	raincoat
**dalu*	*da:lə*	*da:li:*	*dal*	shoulder-blade
**daru-*	*dar(ə)-*	*da:rə-*	*dar-*	to press
**hačï*	*haǰə*	*a:ćə*	---	grandson/cousin
**namur*	*na:mər*	*namur*	*namər*	autumn, fall
**sakï-*	*sa:ġə-*	*sgə-*	*saɣʸ-*	to wait
**bečin*	*be:ǰən*	*mućin*	---	monkey
**beri*	*be:rə ~ bi:rə*	*be:rə*	*bərʸ*	daughter-in-law
**kebi-*	*keβə-*	*ke:yi-*[212]	*kəmʸ-*	to ruminate
**kedü*	*ke:də*	*kədə*	*xəd*	how many
**hodun*	*hɔdən*	*fo:də*	*xɔd*	star
**morïn*	*mɔ:rə*	*morə*	*mərʸ*	horse
**konïn*	*xɔ:nə*	*xonə*	*xɔnʸ*	sheep
**tolï*	*tɔ:lə*	---	*tɔlʸ*	mirror
**tosun*	*tu:sən*	*to:sə*	*tɔs*	fat[213]
**čokï-*	*čɔ:qə-*	*ćugu-*	*čɔkʸ-*	to peck
**dösi*	*tošo*[214]	*do:śə*	*duš*	anvil
**hödün*	*hodən*	*fo:də*	*xudus*	feather
**söni*	*sö:nə*	*sonə*	*sunʸ*	night

Some vowel lengths in Baoan dialects seem to be due to the same phenomenon, e.g. BaoÑ *sa:ġa-* 'to wait', *hə:taŋ* 'star' (related to forms listed above), as well as BaoD

[210] Many of these examples appear in Bolčuluu & Jalsan 1990:62, Čenggeltei 1988:32, 35, 45, 53, Svantesson et al. 2003:113). In some cases the forms with a long vowel will have a variant with short vowel, which will not be mentioned in this list.

[211] This word has a long *a:* in several other languages, which suggests it may belong to the 'ill-spelled secondary length' category discussed earlier. However, the Bargu form *amʊ:* may be either a metathetic length (see below), or it may instead stem from **amusun*.

[212] The vowel length indicated by Mostaert & de Smedt is not confirmed elsewhere.

[213] The loss of CM **toasun* 'dust' may have facilitated this development.

[214] The lowering of the final vowel is unexpected.

wa:r- ‘to grab’ from **barï-*, BaoÑ *na:raŋ* ‘thin’ from CM **narïn*, BaoX *χa:rə-* ‘to return’ from **karï-*.[215]

The following words with (C)VCCV structure usually only have a long vowel in Eastern Yugur (although the unexpected diphthong in MgrH *səulʒə* may count as lengthening)[216]:

CM	E. Yugur	Mongghul	
**maŋgus*	*ma:ŋġəs*	*maŋġusə*	anthropophagous ogress
**ǰančï-*	*ǰa:nǰə-*	*ʒańći-*	to crush
**ǰargu*	*ǰa:rġə*	*ʒarġu*	lawsuit
**gendü*	*ge:ndə*	---	male dog
**mendü*	*mö:ndə/me:ndə*	---	good health
*?*dombur*	*də:mbur*	*dumbur*	hillock
**sonǰï-*	*sə:nǰə-*	*suænʒ́i-* (?=)	to mock
**sölsün*	*sö(:)sən*	*səulʒə*	gall bladder

In a number of cases these lengths correspond with Common Turkic long vowels, as in the case of EYu *ǰa:rġə* ‘lawsuit’ (CT **ya:r-gu* ‘id’). In view of the many other CT words with long vowels that are not reflected in the Mongolic cognate, it seems better to view these as coincidences. Cf. EYu *ġatən* ‘empress’, which corresponds to CT **ka:tun*.

In Eastern Yugur some foreign words of the same structure also have a long vowel. It is not clear what role the quantity in the source language may have played. These lengths are at any rate not confirmed by the other peripheral languages.

**darï*	*da:rə*	gunpowder (ultimately Ir.)
**šabï*	*ša:βə*	pupil (ultimately Skt.)
**šatu*	*ša:tə*	stairs (< Old Uigur <?)

Some less widespread loanwords with similar structure also feature length in Eastern Yugur, including *ga:məs* ‘town’ (of unknown origin)[217], *ǰə:məs* ‘ladle’ and *te:ŋə* ‘deaf’ from Turkic, and *də:rǰə* ‘Indra’s thunderbolt’, *mda:rə* ‘drum’, and *ba:ndə* ‘novice’ from Amdo Tibetan. For the Tibetan loans Mongghul has corresponding forms without vowel lengthening: *dorʒ́ə, ndarə, bandi:*.

Apparently related to this category of lengthening are the ‘morphonological lengths’ in Eastern Yugur, genitive/accusative[218] forms in which the short non-high vowel of the stem was lengthened. Perhaps this does not happen in Mongghul due to

---

[215] Perhaps this explanation also applies to the *a:* of BaoX *ġa:dər* ‘sickle’ < **kaduur* (via a form with shortened **uu*).

[216] Also EYu *ra:ljə-* and MgrH (*a*)*ra:lʒ́i-* ‘to exchange’ (**araljï-*) may owe their long vowels to the word structure. Interestingly this same word lost its middle syllable in Baoan, e.g. BaoÑ *alʒ́a-*. Another case is MgrH *da:ldə-* ‘to trade’ (**kudaldu-*).

[217] Possibly garbled from an older form of Chinese Gānzhōu, the old name of Zhāngyè.

[218] As in the other QG languages, the accusative and genitive are identical in shape in Eastern Yugur.

its genitive in *-nə*, which creates a closed first syllable, an environment less likely to develop a long vowel in Mongghul, as seen above. Eastern Yugur:

Eastern Yugur	CM
ġar 'hand', genitive *ġa:rə*	**gar, *gar-u*
ken 'who', genitive *ke:nə* 'whose'	**ken, *ken-ü*
mal 'livestock', *ma:lə emčə* 'veterinarian'	**mal, *mal-u*
ġal 'fire', genitive *ġa:lə*	**gal, *gal-u*

This might be a recent development. It also occurs in loanwords such as *nag* 'tree', gen. *na:gə* (from Tibetan), *lar* 'speech', gen. *la:rə* (of unknown origin), and even in the second syllable in the Turkic loanword *tu:raġ* 'poplar', gen. *tu:ra:ġə*.

Apart from the words with this structure which have developed a long vowel in Eastern Yugur and/or Mongghul and/or Ñantoq Baoan, there are many similarly structured stems which retained their short vowel in all languages. A small selection: **amïn* 'life', **batu* 'strong', **karï-* 'to return', **tabïn* 'fifty' (apparently to distinguish it from **tabun* 'five'), **tanï-* 'to know', **tarï-* 'to plant', **ǰaru-* 'to use', **ǰarïm* 'half' (unlike **ǰargu* 'lawsuit' from the same stem!), **čerig* 'soldier', **hekin* 'head', **hergi-* 'to turn', **ǰelü* 'tether', **ǰorï-* 'to point', **korïn* 'twenty', **döčin* 'forty', **ökin* 'girl', etc. Most disyllabic words in *-sUn* are not affected either (see the following paragraph).

In the following cases the vowel of the second syllable was lost, so that the lengths may be compensation lengths in a more traditional sense. These words are not very remarkable in that both the loss of final vowels after *-r-* and *Mittelsilbenschwund* are common in Eastern Yugur.

CM	E. Yugur	Mongghul	
**aru*	*a:r*	---	back, behind
**dabusun*	*da:βsən*	*dabsə*	salt
**eri-*	*e:r-*	*yarə-*	to look for
**öri*	*ö:r*	*urə*	debt

Some originally monosyllabic verb stems also developed vowel length. Some of these words may owe their vowel length to the fact that they acquired a final vowel by absorbing the connective vowel, so that they obtained the same (C)VCV structure as the words discussed earlier. Cf. MgrH *ʂdo:lə-* < **ötel-* 'to grow old', *no:rə-* 'to become moist' < **nor-*, apparently through intermediate forms **otoli-*, **nori-*. Similarly, MgrH *ʑo:rə-*, EYu *ǰu:rə-* 'to plane' from **ǰor-*.[219] The length in MgrH *o:sə-* 'to grow' is due to the same development, if related to CM **ös-* 'id'. Alternatively it could be related to EYu *yü:s-* and Dag *əus-*, and stem from CM **eüs-* 'to arise' instead. Another ambiguous form is MgrH *o:li-* (recorded by

[219] These forms and one of the Middle Mongol forms (in the Muqaddimat al-Adab) would suggest a original form **ǰoru-* which has the same structure as the other words with lengthening discussed above.

Mostaert & de Smedt), which form represents both **ol-* 'to obtain' and **bol-* 'to become'. These two verbs merged into a short-vocalic form *olə-* according to other sources. In Eastern Yugur only the former verb has resulted in *ɔ:l-* with a long vowel, whereas *bəl-* remains short.

3.14.6. Length before the element **-sUn*

Vowel lengthening of a non-high vowel preceding the suffix *-sUn* occurs in the central languages, and in Dagur, as well as in QG, in Eastern Yugur, and most frequently in Mongghul. This phenomenon is reminiscent of the lengthening discussed in the previous paragraph, which also involved the lengthening of non-high vowels followed by a high vowel. The two phenomena can not be sharply separated.

Interestingly, most disyllabic formations with *-sUn* are not affected in QG, e.g. **časun* 'snow', **nasun* 'age', **yasun* 'bone', **yosun* 'custom', and **yesün* 'nine'. In the last example **e* had probably been raised to **i* before this lengthening phenomenon arose. **sösün* 'gall bladder' and **tosun* 'grease' did undergo lengthening.

A number of trisyllabic stems may owe the long vowels they display in Eastern Yugur and Mongghul to word structure, without requiring the element *-sUn*. The following similarly structured stems (all with a high first vowel) have length only in Eastern Yugur and Mongghul.

CM	E. Yugur	Mongghul	Dagur	central	
**hünesün*	*henesən*	*funeːʒə*	*xuns*	short V	ashes
**sünesün*	*suneːsən*	*suneːʒə*	*sums*	short V	soul
**hutasun*	*sdaːsən*	*ɻdaːʒə*	---	short V	thread
**sudasun*	*sdaːsən*	*ɻdaːsə*	---	short V	vein
**uŋgasun*	*ŋġwaːsən*	*ŋġuaːsə*	---	short V	wool

The relevance of vowel patterns is underlined by the fact that trisyllables with a non-high first vowel did not develop a long vowel in any language[220], including **arasun* 'skin', **gedesün* 'intestines', **jagasun* 'fish', **kölesün* 'sweat', **olasun* 'hemp'. Indeed, in some stems a development in the other direction can be observed, e.g. EYu *ġadəsən ~ ġasən*, BaoÑ *ġadsəŋ ~ ġarsəŋ* < **gadasun* 'stake'; EYu *gedsən ~ gedesən*, BaoGt *gəsuŋ* < **gedesün* 'bowels'.[221] These reductions are not seen in Monguoric.

The existence of a large number of words ending in a long vowel followed by *-sUn* may have influenced the denominal and opaque formations. The trisyllabic *-sUn* formations are based on diverse stems. Some derive from nominal roots that are known either from Mongolic (e.g. **aduu.sun* 'herd animal') or from Turkic (e.g. **balga.sun* 'wall') while others are not as easily analysed (**arasun* 'skin'). If the

[220] This is confirmed by the very fact that the original high vowels were not lengthened before *-sUn* either.

[221] The reduced shapes are reminiscent of 'western Mongolic', e.g. Kalmuck *gesn*, Moghol *gesən* < **gedesün* 'bowels'.

element preceding the syllable *-sUn* is not recorded independently, it is impossible to determine whether it is in fact a formation with the ending *-sUn*. The fact that the precise original functions of *-sUn* are still debated, makes many words hard to analyse. Another group of words ending in *-sUn* are deverbal formations with *-A-sUn*. The deverbal formations will typically develop a long vowel when the *A* contracts with the final vowel of the stem, or with the connective vowel. E.g. **kala-asun*, **nöke-esün* both 'patch'. Other stem from verbs already ending in a long vowel: **baasun* 'poo', perhaps for **baa-asun* (an inadmissible structure), likewise **kolbaasun* 'connection' for **kolbaa-asun*.

The occurrence of long vowels preceding *-sUn* is erratic, and there even are variant forms within each of the central languages. Some words have long vowels in Buriat and/or Dagur, but short vowels in Khalkha and/or Kalmuck. Other lengths are shared by Eastern Yugur and Mongghul, but not found elsewhere. Buriat and Mongghul agree on **korgaasun* 'sheep dung', whereas Khalkha and Kalmuck have a short vowel in this word. Only Buriat has a long vowel in **balgasun* 'town wall', **sïndasun* 'tendon'. But MgrH *dəra:sə* 'wine' even corresponds to Bur *darhan* with **a > Ø* (CM **darasun*).

The irregular distribution of these lengths makes it nearly impossible to assume their early existence, even on a subgroup level. Only few cases are supported by several subgroups. For instance, EYu *hərči:sən*, MgrH *śʒo:sə* and Bur *üšööhe(n)* 'willow' all seem to suggest a long vowel.[222] Buriat and Dagur point to CM **kïlgaasun* 'coarse hair', but Khalkha, Kalmuck, and Eastern Yugur to **kïlgasun*. The CM reconstructions can not always do justice to the balance of evidence; see the supplement.

Erratic correspondences between the QG languages, Dagur and central Mongolic include the following:

CM	E. Yugur	Mongghul	Dagur	central	
**anǰasun*	*anǰagsən*	*nʒasə*	*anǰa:s*	short V	plough
**burgasun*	---	*burġa:sə*	*baryа:s*	long or short V	willow
**örgesün*	*orgwe:sən*	*rgo:ʒə*	*uryɣis*	long or short V	thorn[223]

Further contradictory modern forms between Dagur and the central languages can be found under **kaïlasun* 'elm tree' and **kekesün* 'spoke'.

3.14.7. Remaining first-syllable vowel lengths in QG languages

Quite a number of further first-syllable vowel lengths can not be explained by secondary lengthening. Most of these lengths occur in one language only.

It seems that non-high vowels tend to be lengthened before *-r-* in Mongghul. Some of these lengths are supported by Eastern Yugur:

---

[222] Eastern Yugur suggests **hičeesün*, Mongghul and Buriat perhaps **hičöesün* < **hičüesün* with complex vowel. The Sino- Mongolian notation *hičesün* in HY would be compatible with the former, as it could be read *hičēsün* just as well.

[223] In view of some central dialect forms this is perhaps really**örgeesün*.

Mongghul	E. Yugur	CM	
mo:r	*mör*[224]	**mör*	road, path
(*xara*) *sa:r*	*sar*	**sar*	certain bird of prey
to:r	*tɔr*	**tor*	net
a:rdaġ	*a(:)rdaġ*	**ardag*	spoilt
ba:rdam	*bardam*	**bardam*	proud
tie:rge (SM)	*teryen*	**tergen*	cart
ya:ra	*ya:ra*	**yara*	ulcer
su:mar (SM)	*su:mal*	**sumal*	bag

Preceding *-r-* even high vowels may be occasionally lengthened, e.g. MgrH *ʒ́iu:rə-* < **ǰïru-* 'to write', *nu:ra:-* ~ *nura:-* < **nura-* 'to collapse'.[225] Lengthening before other consonants occurs less frequently, as in MgrH *ʒ́o:lġo-* (cf. EYu *ǰɔlğɔ-*) < **ǰolga-* 'to meet', and *a:sar* (EYu *hsar*) < **asar* 'pavilion'.

3.14.8. Remaining first-syllable vowel lengths in Dagur

Most of the unconfirmed Dagur vowel lengths (both high and non-high vowels) occur preceding *r* or *n*, suggesting that at least part of them may be secondary developments triggered by this phonetic environment.

Dagur	CM	
a:nəɣ	**anaga*	hunting spot
du:t	**döte*	short (distance)
ga:d	**gada-*	outside[226]
g^{w}ə:n	**gün*	deep
ǰu:r	**ǰüg*	direction
ǰu:r	**ǰür*	roe deer
k^{w}ə:l-	**köl-*	to harness
k^{w}ə:nn^{y}e:	**köŋdei*	hollow
sə:lbə:	**serbee*	fin
sɔ:r	**sur*	thong[227]
ta:wu	**tagu* (?)	jackdaw[228]
xɔ:n	**hon*	year
xu:nu-	cf. **hünir*	to smell

[224] EYu *mór* in Malov's materials (edited by Kotwicz), probably meant to denote **mo:r*.

[225] Lengthening by *-r-* may also be assumed in other languages, e.g. in Ordos *i:r* < **hir* 'knife edge', *i:rge* < **irge* 'wether'. It may also have caused the unexpected Khalkha forms *buurcag* 'bean' and *büürge* 'flea'.

[226] Thus according to Poppe (1967:6). Alternatively this could be a metathetic length from the extended form **gadaa*.

[227] The Dagur variant *saur* given by Zhong could be a secondary development of this long vowel form.

[228] Length is also recorded in Bargu *ta:g* 'id'.

Some superficially similar, but special cases are dealt with in other sections. Dagur *a:n* ‘they’ and *i:n* ‘(s)he’ were perhaps influenced by the other personal pronouns **bi*, **či*, **ba*, **ta*, which were also lengthened in Dagur (see the monosyllables above). The long vowels in Dag *su:* ‘milk’, *xa:yə* ‘bran’, and *x^{w}a:l* ‘tinder’ correspond to long vowels in (some) central Mongolic dialects (see misspelled contraction lengths under 3.14.1.)

Some Dagur lengths may be due to length metathesis, e.g. *xu:l-* < **xu:lə-* < **hulie-* ‘to blow’, *ga:d* < **ga:da* < **gadaa* ‘exterior’ (see 3.14.10.).

Another Dagur length in the first syllable is *t^{w}a:ləɣ* < **tulga* ‘pillar’, which also has a long vowel in Eastern Yugur, but is contradicted by its base, the verb *tɔl-* < **tul-* ‘to support’. Dag *ǰe:bla:-* < **ǰabïla-* ‘to sit cross-legged’ is also unexpected; it cannot merely be a compensation length, given that unaccented vowels are routinely elided in Dagur without the need for compensation.

Further unusual Dagur lengths in the first syllable include *dɔ:l-* ‘(of the weather) to become warm’, perhaps from a verb **dul-*, which could be related to **dulaan* ‘warm’, but is not attested elsewhere, and *nə:ɣ-* ‘to weave’ from **neke-*.

3.14.9. Compensation lengths

In some words vowel lengths seem to have developed as compensation for the loss of a vowel in another syllable. Compensation is an alternative explanation for many of the vowel lengths discussed in the past paragraphs. On the other hand, many apparent compensation lengths can be adequately explained otherwise.

Bolčuluu & Jalsan give several Eastern Yugur words in which the loss of the first vowel is accompanied by lengthening of the accented second vowel: *χwa:r ~ χawar* < **kabar* ‘nose’, *nyɔ:n ~ nion* < **noyan* ‘lord’, *tya:ġ ~ tiyaġ ~ tayaġ* < **tayag* ‘walking cane’, EYu *ɣa:- ~ hia-*, *haya-* < **huya-* ‘to tie up’.

Some of these words also have lengths in Mongghul, such as **noyan* and **huya-*, but this does not necessarily indicate a shared origin, as Mongghul has many lengthened vowels in words of a similar structure, which did not lose any vowels in other syllables. Striking similarities like EYu *na:-*, MgrH *na:-*, BaoÑ *na:-* < **una-* ‘to fall’ become less remarkable in view of the alternative Mongghul form *una:-*.

Compensation in the first syllable for the loss of a vowel in the second may be seen in EYu *da:φsən*, *da:bsən* < **dabusun* ‘salt’, *a:r* < **aru* ‘posterior’, but they may also be due to the lengthening of the non-high vowels discussed in 3.14.5.

Perhaps compensation also plays a role in the length of Dag *g^{w}ə:n* ‘deep’ and *su:* ‘milk’, etc. The forms of **gün* ‘deep’ are mostly short elsewhere, but in Mongghul there are disyllabic forms such as *fugon* which suggests the existence of a CM variants **öken ~ *gü(:)n*. ‘Milk’ also has an old disyllabic variant, cf. Kalmuck *üsn*, Ordos *ᵾsᵾn*, EYu *hsᵾn*, pointing at CM variants **üsün ~ *sü(:)n*. Another similar case may be **onu- ~ *nuu-* ‘to hit’.

Compensatory lengthening may also be due to the loss of a consonant, usually **-b-*. In case of intervocalic **-b-*, this involves contractions of original *VCV* sequences like those in prehistoric Mongolic. Examples include EYu *ǰɔ:lɔŋ* < **ǰoba.laŋ* ‘suffering’, MgrH *de:sə-* < **debis-* ‘to spread’, BaoÑ *te:r-* < **teberi-* ‘to embrace’. EYu *šu:n* ‘bird’ is a contraction of **sïbaun*, but **au* had probably already been contracted before the **-β-* < **-b-* was lost. Puzzlingly three different solutions

were applied in Mongghol to the numerals *de:ren* < **dörben* 'four', *ġura:n* < **gurban* 'three', and *xaran* < **harban* 'ten', the latter without any compensation.

Some of the rare vowel lengths in Mangghuer (recorded by Čenggeltei) have developed as a result of the loss of a consonant, e.g.

Mangghuer	CM	
no:ći-	**nögči-*	to pass
ɕɤǧ ~ ɕɤ:	**čag*	time
ʂa:	**sïra*	yellow

3.14.10. Metathesis of vowel length

In some words vowel lengths seem to be transferred from another syllable. The vowel that was expected to be long on historical grounds appears as short. In fact this is also a type of 'compensation length'.

Among the certain cases of metathesised length we find EYu *toro:-* 'to circle' instead of the expected form **to:rə-* (< CM **toarï-*), BaoÑ *he:lə* 'kind of bird of prey' for **hele:* (< CM **helie*), BaoÑ *ma:ra-*, BaoX *ma:r-* 'to hide' for **mara:-* (< CM **marïa-*).

Dag *ga:d* 'outside', if from **gada:* < CM **gadaa*, could belong in this group as well. However, Poppe (1967:6) compares it to the base **gada* (which he reconstructs **ġa:da*) rather than to the derivate **gadaa*.

One of the more complicated cases is Dag *xu:l-* 'to blow', which should be related to CM **hülie-*. This etymon has an unexpected short final vowel in all peripheral languages, a diphthong in Dongxiang, and a long vowel in the first syllable in Mongghul and Eastern Yugur. Eastern Yugur and Mongghul further have initial [p^h-], showing that paralinguistic factors have played a role.

The verb **dolïa-* 'to lick' lacks the expected contraction vowel in the QG languages, where all modern forms seem to derive from a form **do:lï-* or **do:l-* 'to lick' instead of expected **dolo:-*. However, in this case the length may also have arisen after the word developed the syllable structure **dolï-* (see 3.14.5). This would leave open the question why its second syllable was shortened.

BaoX *ġa:dər* 'sickle' (ultimately from **kaduur*) is another example where the length may be due to syllable structure, seeing that other Baoan dialects feature the form *ġadər*. Length metathesis in Baoan may go hand in hand with the tendency to reduce contraction lengths in the non-first syllables.

In the peripheral languages there are only isolated cases of vowel length metathesis. In Bargu, long vowels before the suffix **-sUn* are often transferred to the second syllable.

Bargu	CM	
baxa: (for expected **ba:xa*)	**baasun*	excrement
buxə: (for expected **bu:xə*)	**böesün*	louse
nɔxɔ: (for expected **nɔ:xɔ*)	**noasun*	wool

Analogy is probably the reason why today most words containing the suffix *-sUn* now end in a final long vowel in Bargu, even if the first syllable vowel was not long:

naxa: (for expected **naxa*)	**nasun*	age
tɔxɔ: (for expected **tɔxɔ*)	**tosun*	oil
ʊxʊ: (for expected **ʊxa*)	**usun*	water

The same analogy may have affected Brg *buxə:* < **büse* 'belt', unless this is a genuine length (cf. also the QG cognates MgrH *puse:* and EYu *pəsəi*).

3.14.11. Lengthening of non-first vowels in Dagur

Apart from the monosyllabic stems and some cases which are also long either in central Mongolic, or in Mongghul and/or Eastern Yugur, most other unexpected vowel lengths in Dagur occur in non-first syllables. These lengths appear to be totally unrelated to QG lengthening, and they are mostly unexplained. Perhaps most involve morphological reinterpretation rather than phonetic developments.

Although there is no explanation for the secondary lengthening of these vowels, it is problematic to conclude, on the basis of Dagur alone, that they are therefore 'primary' lengths (as suggested in Tömörtogoo 1992). The corresponding vowels are short in the QG languages Eastern Yugur, Mongghul, and Ñantoq Baoan, which even tend to place the accent on the final syllable. Several groups of similarly-structured words can be distinguished. They will be discussed below.

3.14.11.1. Lengthening of the final vowel of verbs

A number of, mostly disyllabic, verbs end in *a:* and *ə:* where the vowel was expected to be lost, e.g. *xumpa:-* (**humba-*) 'to swim' and *xuya:-* (**huya-*) 'to tie up', *təmče:-* (**temeče-*) 'to strive'. The occasional agreement with lengths in Mongghul (as in the three verbs just mentioned) seems to be coincidental.[229] If they do represent old lengths, the short vowels in central Mongolic and Eastern Yugur need an explanation.

While lengthening of final vowels is quite common in Mongghul, it is hard to explain in Dagur, unless some kind of metanalysis is assumed, whereby original causative formations in *-A* supplanted the original verbs, or some frequent inflected forms such as the converb *-AAr (*-Ad)* may have caused the (historically) incorrect reinterpretation of the stem.

An explanation in that direction is supported by the fact that the original vowel was not simply lengthened, whichever quality it had. Rather a non-high vowel element *A* was added to the original stem. Although *A* is also the shape of a causative ending, the change in shape did not affect the meaning of these verbs.[230] Thus Dagur has *əlwə:-* < **elgü-* 'to hang', *ərwə:-* < **ergü-* 'to lift' rather than forms in long *-u:*.

---

[229] **huya-* 'to tie up' also developed a long final vowel in Buriat.

[230] Nevertheless it may be relevant that most of the affected verbs are transitive.

Verbs in **i*/**ï* have Dagur *e:* (/*ie*/, [^{y}*e:*]) instead of expected /*i*/ [^{y}] in verbs such as *kumr^{y}e:-* < **kömeri-* 'to overturn' and *təur^{y}e:-* < **teberi-* 'to embrace', *karče:-* < **kabčï-* 'to squeeze', *kɔmb^{y}e:-* < **kumï-* 'to fold', *xɔl^{y}e:-* < **kolï-* 'to mix', *tɔbk^{y}e:-* < **tobkï-* 'to stitch coarsely'.[231] Here simple lengthening of the final vowel would have resulted in Dagur *i:*.

The age of this development is unknown, but the development of a (Dagur and central Mongolic) form **bulïa-* 'to snatch' alongside **bulï-* may be an early example.

For verbs ending in the 'long-vocalic suffixes' *-lA, -gA, -lGA* see 3.14.12. below.

3.14.11.2. Lengthening of the last vowel of disyllabic nouns

This type of lengthening typically affects non-high vowels. Surprisingly often these Dagur lengths correspond to short vowels in Mongghul, although Mongghul otherwise has many vowel lengths in disyllabic words of similar structure (see 3.14.4.).

Dagur	cf	CM	
bəryə:n	EYu *be:rgə*	**berigen*	sister-in-law
guɣa:n		?**göen*	shallow
gura:n	EYu *ġura*	**gura*	roebuck
daŋga:l	MgrH *daŋġʊl*	**daŋgal*	clod
x^{w}aima:r	MgrH *xʊimor*	**koïmar*	seat of honour; Mgr: floor

This category also includes the following adjectives, again contradicted by Mongghul (forms from Junast here). In some cases the long vowels are also contradicted by Zhong's Dagur data.

Dagur	Mongghul	CM	
niŋgə:n	*nəŋgen*	**nimgen*	thin, fine
šiŋgə:n	*śiŋgen*	**siŋgen*	thin (fluid)
xuŋgə:n	*koŋgon*	**köŋgen*	light (weight)
yauɣa:n	---	**yabugan*	on foot
ǰəulə:n	*ʒo:lon*	**ǰöelen*	soft

All these lengths correspond to short vowels in Mongghul, even though Mongghul has a tendency to lengthen second syllable non-high vowels followed by a nasal or liquid. However, the reverse situation is seen in Dag *urkun* < **ödken* 'thick', as opposed to MgrH *ɻdoġo:n* 'id'.

[231] Dag *kəkre:-* 'to belch' and *təmče:-* 'to compete' suggest CM forms in *-i* rather than the forms **kekere-* and **temeče-* supported by other languages. Conversely, Dagur *nərčə:-* 'to hatch' suggests **nebčee-* rather than the usual **nebčie-*.

There may be a connection between the above unexpected Dagur lengths and the lengthening of the vowel before the diminutive *-kAn* in Dagur adjectives.[232] Remarkably, some of the adjectives have short vowels in their final syllable although they contained CM double or complex vowels.[233]

Dagur	CM	
bələn - bələːkən	**belen*	(quite) ready
bulun - buluːkun	**bülien*	(quite) luke-warm
narin - nariːkən	**narïn*	(quite) fine[234]
šɔlun > šɔluːkun	**sïluun*	honest > nimble
xairən - xairaːkən	**kaïran*	(quite) pitiful

3.14.11.3. Lengthening of last syllable vowels before *-r*

Lengthening is found in several Dagur words with CM **-r* or secondary *-r* from **-g* or **-d*. The vowel is typically non-high. Beyond the fact that *r* is known to often lengthen the preceding vowel, there is no explanation for this phenomenon. Cf. lengthened monosyllables in *-r* in 3.14.8.

Dagur	Mongghul	CM	
alaːr	*alaġ*	**alag*	variegated
baɣəlǰaːr	*baġalʒaġ*	**bakalǰag*	wrist
bədəːr	---	**beder*	spot, freckle
bɔdɔːr	*budəġ*	**budug*	dye
bulaːr	*bulaġ*	**bulag*	(water) spring
kaltaːr	---	**kaltag* or **kaltar*	name of a horse's coat colour
kɔnɔːr	---	**konag*	24-hour period
kɔŋgɔːr	---	**koŋgar*	light brown

These examples raise the question why this type of lengthening is not more universally applied (did *kɔdir* < **kudug* 'water well' and *suǰir* < **süǰüg* 'religion' remain short due to the high vowel?). Another problem is that many of the words in question could be (re-)borrowings from Manchu, which has *bederi* 'spot', *kaltara* 'brown with a white nose', *konggoro* 'brown'. This also applies to the adjectives above, many of which also occur in Manchu, e.g. *bulukan* 'lukewarm', *hairakan*

[232] In Mongghul the suffix **-kAn* has developed a long vowel, compare Mongghul *bulenxaːn*, as opposed to Dagur *buluːkun*.

[233] Puzzlingly, Enkhbat (1988:235) mentions the stems as *buluːn, nariːn, šɔluːn*, while his dictionaries, as well as Zhong's materials, report the same words with a short vowel in the unexpanded stem. Did he 'invent' these word shapes to explain the morphology on that given page? Alternatively, the lengthening is due to analogy with other words mentioned by Enkhbat, in which the length is due to contraction, such as *xuluːkun* 'quite many', *xaluːkun* 'quite hot', from **hüleü* and **kalaun*, respectively.

[234] In view of Kh *nariyn* 'fine', the *-i-* of Dag *narin* could also be viewed as a shortening.

'pitiful'. However, to assume reborrowing from Manchu does not ultimately lead to a solution, as many of the vowel lengths of Manchu (as recorded in Sibe) are themselves unexplained.[235] Moreover, some of these words also display, partly erratic, vowel lengths in the neighbouring North Tungusic languages Solon and Oročen. In all these cases it is difficult to establish the direction of borrowing[236], and at any rate none of the languages offers an obvious explanation for the long vowels.

Lengthening of CM high vowels is much rarer in Dagur. Dag *anʸe:r* in fact suggests a CM form **anïar* rather than **anïr*.

Dagur	CM	
anʸe:r	**anïr*	sound
xuli:r	**ölir (?*hölir*)	wild apple

3.14.11.4. Various other Dagur vowel lengths

Several other Dagur words have unexpected vowel lengths. Lengthening of the last vowel of trisyllabic or longer nouns, either word-finally or preceding a sonorant is unexpected because of the Dagur accent on the first syllable.

Dagur	E. Yugur	CM	
karəɣna:	*qarğanə*	**karagana*	(plant name)[237]
ərɣʷlə:	*örlö(:)*	**öreele*	hobble
naučɔ: ~ nauč	*nağačə*	**nagaču*	maternal uncle

The long vowel in *naučɔ:* may be not be a phonetic development. It may be a vocative or caritative form, cf. Dag *əwə:* 'mother', *əkə:* 'older sister' and *aka:* 'older brother', which apparently are vocative forms to the regular developments *əɣ* < **eke*, *əkč* < **egeči*, and *aɣ* < **aka*.

Lengthening of the second vowel of trisyllabic nouns, comparable to the group of *ala:r*, *bədə:r*, discussed above, but with other sonorants instead of *r*. These cases are also unexpected because of initial stress.

Dagur	CM	
čakɔ:lʸ	**čakulaï*	seagull
nida:nʸ	**nïdanï*	last year
kurʸe:n	**kürene*	weasel ('vole' in Dagur)
təmə:nʸ	**təbənə*	large needle

[235] There are other cases where Dagur has vowel lengths in Mongolic words reborrowed from Manchu, e.g. Dag *xada:l* 'horse's bit' and *durə:ŋgʸ* (Zhong), *durə:ŋgi:* (Enkhbat) 'stirrup' from Manchu *hadala*, *durenggi* (cf. CM **kadaar*, **döree*). But these vowel lengths are not documented in Manchu. Perhaps they were just perceived by Dagur as not reduced, and therefore associated with long vowels.

[236] Solon also has the development *-r* from **-g*.

[237] Cf. also Dag *tuɣna:* 'lead (the metal)', apparently related to **tuulga*.

Again there are equivalents in Manchu, which cannot explain the vowel lengths, e.g. *kurene* 'weasel', *temene* 'large needle'.[238]

Dag *səlmi:* 'sword' from **seleme* is inexplicable unless it is originally a genitive which was morphologically redefined as an uninflected root. Dag *tulma:* 'bucket' (CM **tulum*, with no final vowel at all) may have been a reflexive possessive form reinterpreted as the stem. These explanations are rather *ad hoc*, but there is no purely phonetic development to explain these forms. Similar reinterpretations may underlie other cases where a Dagur long vowel corresponds to *Ø* in other languages, such as the verbs mentioned under 3.14.12.

In closing, there are many peculiar vowel lengths in Dagur: In view of the chaotic correlation with vowel length in QG and elsewhere, it is not generally advisable to make these poorly supported vowel lengths part of the CM reconstruction forms. Perhaps the most striking aspect of the Dagur vowel lengths is that so many occur in unaccented syllables. This may in fact indicate that accent placement was much more variable in earlier stages of development.

3.14.12. Long vowels in suffixes

The vowels of some suffixes have apparently been lengthened in Dagur and Mongghul. Not the same suffixes are lengthened in each language, so that it is impossible to reconstruct these suffixes as being long in CM. There is no explanation, other than the possibilities of a) metanalysis of derived or inflected forms, and b) the influence of neighbouring languages where the same suffixes exist.

As even contraction lengths can be found shortened in modern languages, it can not be excluded that some of the long vowel suffixes represent the original situation, and that their short counterparts elsewhere are the innovation.

Disagreements with respect to vowel length can be seen in several inflectional and derivational suffixes. The ablative suffix (probably from CM **-AAsA*, in spite of Written Mongol *-AčA*) has become *-A:s* in Dagur, but *-sA* in Eastern Yugur and *-sa* in Mongghul. The comitative **-lUA* > **-lA:* has become MgrH *-la*, and the ordinal suffix **-dUAr* has been shortened into MgrH *-dar*.

A more thorough investigation will be required, taking into account the many disagreements between various descriptions, especially in the case of Dagur. In the following only a selection of suffixes with lengthened vowel will be presented.

The denominal verb suffix *-lA*

In Dagur the vowel of the denominal verb suffix **-lA* is lengthened in a large number of derivations. However, another large group of *-lA* derivations has the expected reflex *-l*. It is not clear why some words developed the long vowel, and others did not. As there appears to be little free variation, it seems to be lexically determined.[239]

---

[238] See Kara 1985 on the Dagur vowel lengths in Manchu loanwords.

[239] One case where both the short and the long form occur is *wəyil-* and *wəylə:-*, both from CM **üye* 'joint', but with diverged meanings.

In Eastern Yugur, Mongghul, and Ñantoq Baoan, the suffix **-lA* is normally short, in spite of the fact that final non-high vowels are oftened lengthened in these languages.[240]

The verbalizer *-lA* is an 'Altaic' suffix also occurring widely in Tungusic, and its vowel is also long in Solon and Orochen, two Northern Tungusic languages neighbouring Dagur. Whatever the cause of the length in these languages, it is quite possible that these languages influenced Dagur in this respect (since Tungusic influence on Dagur is considerable).[241]

Most verbs in *-lA* are straightforward verbalizer-derivates in which the nominal stem from which it is derived can be easily recognized. For the (usually predictable) meanings of most verbs I refer to the dictionaries. The long-vowel version of the suffix appears in: *aula:-* (from **aba* 'hunt'), *ča:sla:-* (from **čaasun* 'paper'), *dəɣə:lə:-* (from **degee* 'hook'), *dəllə:-* (from **dere* 'pillow'), *gujilə:-* (**güjir* 'difficult'), *mullə:- ~ murlə:-* (**mörü* 'shoulder'), *ləllə:-* 'to call' (**nere* 'name'), *nuɣ^{w}lə:-* (**nüken* 'hole'), *t^{w}ə:lə:-* (**töe* 'handspan'), *usla:-* (**usun* 'water'), *x^{w}arəmla:-* (**kormaï* 'hem'), *x^{w}arkla:-* (**huraka* 'lasso').

The short-vowel version of the suffix: *čə:rəl- ~ čə:l-* (**čeer* 'prohibition'), *daul-* (**daun* 'song'), *dabkul-* (**dabkur* 'layer'), *də:səl-* (**deesün* 'rope'), *dɔmul-* (**dom* 'magic'), *najil-* (**najïr* 'summer'), *wə:dəl-* (**öede* 'above'), *kəryəl-* (**kereg* 'need'), *nasəl-* (**nasun* 'age'), *sujil-* (**süjüg* 'religion'), *t^{w}a:l-* (**toan* 'number').

These examples show that both versions of the suffix can be attached to native noun stems, and that there is no obvious phonetic reason that could explain the choice of suffix variant. As word-final short vowels are normally reduced or lost, the second set of derivations show the expected development.

The deverbal verb suffix *-lA*

Some verbs which at first sight have the same structure as the above formations with *-lA*, in fact contain different suffixes. There is an intensive/iterative (deverbal verb) suffix **-lA* which survives in Dag *nɔɣ^{w}la:-* 'to pound continuously', from *nɔɣ^{w}-* < CM **nïku-* 'to knead', *murɣ^{w}lə:-* 'to butt (of cattle)' from *murɣ^{w}-* < **mörgü-* 'to bow; to butt'. This suffix, too, has a short vowel elsewhere: EYu *qudqula-* 'to stir' from **kudku-*, MgrH *nuǧla-* 'to knead' from **nïku-*, BaoÑ *ʒabćila-* 'to chop up' from **čabčï-*.

The transitive marker *-l*

Even more peculiar are the verbs formed with the transitive marker **-l* (counterpart of intransitive **-rA*), which did not even contain a vowel originally. Apparently this suffix was associated with and influenced by the verbalizer *-lA* and/or other transitive (causative) verb formations ending in the vowel *A*. Metanalysis of suffixed forms (e.g. the converb *-AAr* or the agent noun *-AAči*) may also have contributed to

[240] As usual, there are some exceptions, e.g. BaoÑ *ɤirla:-* 'to become yellow'< **sïra-la-*.

[241] The Turkic language Yakut, also has a long vowel in this suffix, which seems to be confirmed by the spellings in Kashghari's 11th century dictionary (which could be a matter of spelling).

this unexpected development. Examples include *sətlə:*- (**setel*-) 'to tear', *xaulɔ:*- (**kagal*-) 'to break', *xɔɣ^{w}lɔ:*- (**kugul*-) 'to break'. However, not all *-l* formations have been lengthened, cf. *xaul*- 'to skin' (**kaul*-). Compare also the morphologically opaque verbs in *-l* showing the same long vowels such as Dag *kutlə:*- from **kütel*- 'to lead by the hand', *sɔndla:*- < **sundula*- 'to ride with a passenger or luggage'.

The denominal verb suffix *-rA*

The development of the denominal verb suffix **-rA*, mainly a verbalizer of adjectives, shows the exact opposite development of **-lA*: its vowel is lengthened in Mongghul, and reduced or lost in Dagur. The corresponding forms in Eastern Yugur do not present a clear picture. Junast (1981:86) records the suffix as *-rAA*, but Bolčuluu & Jalsan (1991:383) as *-rA*. Cf. MgrH *ġalʒ́u:ra:*-, Dag *galǰɔ:r*- < **galǰaura*- 'to go mad', MgrH *kuidəre:*-, Dag *kuitur*- < **küitere*- 'to become cold'.

Dag *šarəl*- ~ *šallə*- 'to become yellow', and *xarəl*- ~ *xallə*- 'to become black' are from CM **sïra-la*- and **kara-la*-, but these are dissimilated forms of earlier **sïra-ra*- and **kara-ra*-, which are derived by means of the suffix **-rA*. The vowel of this suffix is not lengthened in Dagur, which is demonstrated by the development of other derivations such as **köke-re*- 'to become blue' (Dagur *kukur*-).

In view of the numerous other verbs with lengthened final vowel in Mongghul, the lengthening of verbs in *-rA* may not seem to be surprising. However, this lengthening is striking when compared with *-lA* formations which nearly always remained short-vocalic in Mongghul.

The intransitive marker *-rA*

The other suffix **-rA*, intransitive counterpart of **-lA*, is also lengthened in Mongghul, as illustrated by verbs like MgrH *xġara:*-, Dag *xaɣər*- < **kagara*- 'to burst', MgrH *taġsəra:*-, Dag *tasər*- < **tasura*- 'to break (intr.)', MgrH *xaura:*-, Dag *xaur*- < **kaura*- 'to peel off (intr.)'. In Eastern Yugur the suffix has the form *-rA* according to Bolčuluu & Jalsan, but it occasionally appears with long vowel in Junast's materials.

The denominal verb suffix *-dA*

Like **-lA*, the transitive verbalizer **-dA* has two sets of reflexes in Dagur, with long and short vowel, without obvious reason. In Mongghul and Eastern Yugur, **-dA* has a short vowel. Dagur examples with long vowel include *čɔlɔ:da:*- 'to throw stones' from **čïlaun* 'stone', *xaurda:*- 'to file' from **kauraï* 'file (the tool)'; examples with short vowel include *kirə:d*- 'to saw' from **kiröe* 'saw', *daud*- 'to read' from **daun* 'voice'.

The adjective modifier *-bIr/-bUr*

The adjective modifier **-bIr/-bUr* has a long vowel in Mongghul, but a short one in Ñantoq Baoan and Eastern Yugur, e.g. MgrH *fulaanwu:r*, BaoÑ *fulawər*, EYu *łaiβər*

'reddish', from **hulaa.bur* or **hulaa.bïr*.[242] The short vowel is confirmed by Dagur, assuming that the functional equivalent *-lbin/-rbin* in *xula:lbin ~ xula:rbin* is etymologically related. The long vowel in Mongghul *-wu:r* seems to be doumented only in Li Keyu's materials, cf. *ćiġa:vur* 'whitish' (Činggeltei 1988:356).

The diminutive **-kAn*

The diminutive **-kAn*, found on numerals and adjectives, is also long in Mongghul, but short in Ñantoq Baoan, Eastern Yugur and in Dagur. Mongghul examples include *kundunxa:n* 'rather heavy' from **kündü* 'heavy', *xaranxa:n* 'only ten' from **harban* 'ten'. As mentioned above (3.14.11.2.), the vowel *preceding* the suffix **-kAn* may be lengthened in Dagur, as in *bələ:kən* 'rather ready' < **belen* 'ready'.

The causative suffixes *-lgA, -gA*

The causative suffixes **-lgA* and **-gA* have usually become Dagur *-lɣA:*, *-lkA:*, *-gA:*, *-kA:* with long vowel. Examples include Dag *ailɣa:-* (from **ayu-* 'to fear'), *darəlɣa:-* (from **daru-* 'to press'), *du:rɣə:-* (from **düür-* 'to fill up'), *garɣa:-* (from **gar-* 'to exit'), *saulka:-* (from **sau-* 'to sit'), *sɔrɣa:-* (from **sur-* 'to learn'), *daurka:-* (from **daus-* 'to end'), *xəirkə:-* (from **keis-* 'to be blown away'). Perhaps these causative formations were lengthened under the influence of other causative verbs with a long final vowel, originally formed with *-A*, such as *kata:-* < **kataa-* 'to harden', causative to *kat-* <**kata-* 'to become hard', *dərtə:-* < **debtee-* 'to soak (tr)', from *dərt-* < **debte-* 'to soak (intr.)'. That the lengthening is not a phonetic development is shown by verb stems ending in the syllable *gA*, which regularly lost their final vowel, e.g. Dag *ədəɣ-* < **edege-* 'to recover', *itəɣ-* < **itege-* 'to believe', *yaləɣ-* < **ïlga-* 'to choose', *ǰaləɣ-* < **ǰalga-* 'to join', *ǰarəɣ-* < **ǰïrga-* 'to rejoice'.

The lengthening of the vowel of **-lgA* and **-gA* is not confirmed by Eastern Yugur and Mongghul, where these suffixes typically remain short-vocalic. Mongghul only has a long vowel in a handful of older causatives, e.g. *kurge:-* < **kürge-* 'to bring, deliver', *surġa:-* < **surga-* 'to teach'. These are no longer felt to belong to the verbs *kur-* 'to reach' and *suru-* 'to learn'; Mongghul has created new causatives *kurġa-* 'to cause to reach', *suruġa-* 'to make sb. learn' with a closer semantic relation to the base verb.

The deverbal noun suffix *-l*

In Dagur the deverbal noun suffix *-l* is now often preceded by a long vowel which was not present in the verb stem, e.g. *itɣə:l* < **itege-l* 'belief', *ǰarɣa:l* 'smallpox' (Ersatzwort) < **ǰïrga-l* 'happiness', *ǰɔɣɔ:l* < **ǰoba-l* 'problem', *mədə:l* < **mede-l* 'knowledge'. At first sight these derivations suggest that the suffix somehow triggered the lengthening of the stem vowel. However, *xɔrɣ^{y}e:l* (rather than **xɔrɣi:l*) < **horgï-l* 'whorl' suggests that the shape of the suffix itself has been redefined as

[242] The *-n* of the stem **hulaan* was unexpectedly retained before *-bUr* in Mongghul. The Eastern Yugur form was probably influenced by the related verb *ɬai-* < **hulaï-* 'to become red'.

-Al or *-AAl*, so that *xɔrɣ^{y}e:l* represents **horgï-al*. The shape *-AAl* is suggested by derivations from verbs ending in a consonant, such as *ɔrsɔ:l* 'stream' from *ɔrs-* < **urus-* 'to flow'. Perhaps this redefinition was also inspired by formations that do derive from stems in a long vowel, such as *xasɔ:l* < **hasau-l* 'question', *xirə:l* < **hiröe-l* 'blessing', although these are not frequent. As in the case of *-lA*, there are also forms with the expected phonetic development, i.e. a short vowel: *šadəl* < **čïda-l* 'ability', *dɔsul* < **dusu-l* 'drop', *təsəl* < **tes-ü-l* 'patience'.

In other languages this suffix tends to end in *-l* only. There are isolated parallels in central Mongolic, e.g. Khalkha *magtaal* 'praise' and *surgal ~ surgaal* 'doctrine', which should be from **magta-l* and **surga-l*. In Mongghul a long vowel may occur when the verb stem itself had developed a long vowel.

3.14.13. Dagur lengths due to Tungusic cognates

Several Dagur lengths occurs in Mongolic words reborrowed from neighbouring Tungusic languages. Potentially a large number of the Dagur vowel lengths discussed in the previous paragraphs may involve reborrowing from Tungusic, given the sizeable vocabulary Dagur shares with its neighbours. In some instances the borrowing has led to doublets, typically with a different semantic nuance, e.g. Dag *kara:* 'black (of a horse's coat)' is from Oročen, while native **kara* resulted in Dag *xar* 'black (in general)'. Dag *širɣa:* 'yellow (of a dog's coat)' is probably from Manchu, whereas native **sïrga* resulted in Dag *šarəɣ* 'straw-coloured with light mane (horse's coat)'. Dag *sula:* 'idle' is from Manchu, whereas native **sula* resulted in Dag *s^{w}al* 'loose; free'.[243]

3.15. Turkic evidence for 'primary' vowel length?

As seen in the previous paragraphs, the internal Mongolic evidence for 'primary' vowel length, although at first sight abundant, is quite ambiguous. When the numerous Dagur, Eastern Yugur, and Mongghul words which feature interesting vowel lengths are investigated more closely, it appears that many such lengths are only supported by a single language, and rarely by two or three of the peripheral languages. In many cases these languages disagree on the quantity of vowels, both in endemic words such as **berigen* 'sister-in-law', and in words with CT cognates such as **kekire-* 'to belch'. Therefore, the modern Mongolic evidence generally does not suffice to reconstruct these lengths for the ancestral language.

In spite of the contradictory evidence, unexplained vowel lengths found in the individual languages might be remnants of an earlier (Proto Mongolic) situation. As vowel lengths can often be confidently reconstructed in Turkic cognates, one would expect these cognates to provide some additional evidence for lengths in Mongolic.

[243] The vowel lengths of Manchu are not indicated in spelling, but they survive in its modern dialects including Sibe.

In the following the potential correspondences between Turkic and Mongolic long vowels will be explored. The usual correspondences between Mongolic and Turkic with regard to vowel quantity are as follows:

1) CM simple vowel corresponding to a short vowel in Common Turkic
(e.g. CM **kara* - CT **kara* 'black')

2) CM simple vowel corresponding to a long vowel in Common Turkic
(e.g. CM **sal* - CT **sa:l* 'raft'; CM **ǰarïm* - CT **ya:rïm* 'half')

3) CM double vowel corresponding to a *VCV* (or *VC*) sequence in Common Turkic
(e.g. CM **ǰaïdag* - CT **yapïdak* 'without saddle'; CM **saarï* - CT **sagrï* 'rump')

A fourth common correspondence should be mentioned here as well:

4) CM simple vowel corresponding to *Ø* in Common Turkic
(e.g. CM **bora* - CT **boz* 'grey'; CM **köke* - CT **kö:k* 'blue')

Similarities in vowel length between modern Mongolic languages and Turkic[244] are very rare. Even in the more striking cases there tend to be problems with the comparison, either the semantic development makes it difficult to establish the identicity of a Mongolic word and its Turkic counterpart, or the length itself is only supported by a single language, or can be explained as a secondary development.

CM **dalu* 'shoulderblade' (whose **a* resulted in vowel length in Eastern Yugur and Mongghul) could be related to CT **ya:l* 'mane', but the Mongolic length can be explained as secondary lengthening due to word structure.[245]

CM **ǰargu* 'lawsuit' (with long vowel in Eastern Yugur) is obviously related to CT **ya:rgu* 'lawsuit', derived from the verb **ya:r-* 'to split'. However, CM **ǰarïm* 'half', corresponding to CT **ya:rïm* 'half', and derived from the same Turkic verb, lacks the vowel length in Mongolic. CM **yara* 'wound', possibly also related to the same verb, looks more convincing in that it has a long vowel in Eastern Yugur, Mongghul, and Ñantoq Baoan, but it is contradicted by Dag *yar*. Moreover, the reconstruction of length in Turkic is also problematic, as Turkmen has a short-vocalic form *yara* suggesting CT **yara* without length, and perhaps unrelated to the verb.[246] CM **ǰolga-* 'to meet' (with long vowel in Mongghul) is

[244] There are several types of evidence for the original presence of vowel length in a Turkic stem. Long vowels as such survive in Turkmen, Yakut, and Khalaj. Lenition of the consonant following the vowel in Oguz languages is another indication. Many short vowels can be recognised in Tuva and Western Yugur by means of the accompanying pharyngalisation and preaspiration, respectively.

[245] In CT there is also a front-vocalic form **yɛ:l* 'mane' which corresponds to CM **del* 'id'.

[246] Moreover, *yara* seems to be unattested in earliest Turkic, and is morphologically problematic. The original meaning may have been 'sore, ulcer', which then later developed into 'wound' under the influence of the verb **ya:r-*. Unlike **ǰargu* and **ǰarïm*, CM **yara* has an initial *y-* in the QG languages. This suggests that, if all three words ultimately derive from the same verb, these correspondences belong to different phases of Turkic-Mongolic contact.

clearly related to CT **yo:luk-* 'to meet'. However, this length is only confirmed by a single Mongolic language.

Such problematic cases would not be relevant if there were hundreds of clear correspondences. In fact in the majority of cases there is no conclusive evidence in Mongolic for vowel lengths that are already known to exist in cognate Turkic words. Most of the etyma listed below are present in Eastern Yugur and/or Mongghul and/or Dagur, and normally occur without length.

The small selection below demonstrates how even an initially striking correspondence need not be very convincing. Not only the contradiction between languages (and between descriptions of the same language) damage the credibility of 'primary' vowel lengths in Mongolic, but also the fact that there are alternative explanations for several of them.

CT	CM		length supported by	contradicted by
**sa:k-*	**sakï-*	to wait	EYu *sa:ġə-*, BaoÑ *sa:ġa-*	Dag, Mgr
**ke:kïr-*	**kekire-*	to belch	EYu *ge:gəre-*	Dag, Mgr
**ye:l*	**del*	mane	Dag *də:lbur*	EYu, Mgr[247]
**so:l*	**solagaï*	left	Dag *saulɣui*	EYu, Mgr[248]
**to:r* (?)	**tor*	net	MgrH *to:r*	EYu[249]
**yo:luk-*	**ǰolga-*	to meet	MgrH *ʒ́o:lġo-*	EYu, BaoÑ
**kö:l- (?*kü:l-)*	**köli-*	to harness	Dag *k^{w}ə:l-* (as if < **kö:le-*)	EYu, Mgr

In other cases Mongolic lengths correspond to short vowels in Turkic. The vowel length suggested by MgrH *bu:rge* 'flea' is different in that it is supported by Kh *büüreg*[250], but in Turkic it also has a short vowel. In the following cases the Mongolic length has arisen due to word structure (see 3.14.5.).

CT	CM		length suggested by	contradicted by
**hatï*	**ačï-*	grandchild	MgrH *a:ćə*	EYu
**koñ*	**konïn*	sheep	EYu *χɔ:nə*	Dag, Mgr
**ya(g)ku*	**daku*	raincoat	MgrH *da:xu*	Dag

[247] It is not unlikely that the first syllable of the Dagur word represents CM **del*. The etymology of the second syllable is unknown to me.

[248] This Dagur form is only recorded by Zhong, elsewhere forms with *-ɔ-*. The *-au-* may, but need not be represent original length, a parallel case with the same alternation is Dag *saur* ~ *sɔ:r* 'thong' (also an unexpected length).

[249] The Mongolic length seems to be confirmed by Written Mongol *toor*, but this double spelling has proven unrelated to vowel length in other cases. The Turkic length is not entirely certain either, as Turkmen has *tor* with short vowel; Khalaj *tūor* could be a borrowing from Persian, where this word is also found.

[250] Kh *buurcag* is neither confirmed by the peripheral languages, nor by CT.

In yet other cases established Turkic vowel lengths correspond with short vowels in all of Mongolic:

CT	CM	
**a:la*	**alag*	multi-coloured[251]
**a:rt-*	**a(r)čï-*	to load
**ba:gla-*	**bagla-*	to tie up
**ba:y*	**bayan*	rich
**sa:l*	**sal*	raft
**sa:na-*	**sana-*	T: to count; M: to think
**e:r*	**ere*	man (if at all related)
**bo:nčok*	**mončag*	bead
**(h)ö:rmek*	**örmege*	coarse fabric
**kö:k*	**köke*	blue
**sö:k-*	**sögee-*	to scold
**bu:ka*	**buka*	bull
**su:n-*	**sun-*	to extend
**kü:č*	**küčün*	strength

All in all, the correspondences between Turkic and Mongolic are erratic at best. The abovementioned examples suggest that, if CT lengths have ever regularly corresponded to long vowels in CM or PM, they have now been lost in both the central and the peripheral languages. The non-contraction vowel lengths that do occur in the peripheral Mongolic languages usually correspond to CT short vowels (if a Turkic cognate exists).

[251] The Dagur form *ala:r* has an unexplained vowel length in the second syllable. As Khalaj has *h-* in this word, it is not certain that the Turkic and Mongolic words are even related.