



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).

1. INTRODUCTION

1.1. The Qīnghǎi-Gānsù languages and reconstruction

The Mongolic languages spoken in Qīnghǎi and Gānsù provinces in northern China have struck researchers by their aberrant developments, compared to the central Mongolic languages. Some unique features were recognised as being in part due to contacts with neighbouring languages, while others were clearly ancient features, reminiscent of Middle Mongol. In the past decades the Mongolic languages of Qīnghǎi and Gānsù have become better known. Descriptions are available for all known Qīnghǎi-Gānsù languages¹, as well as numerous publications on various diachronic matters. It is now possible to study these languages as a group from a comparative historical viewpoint.

This book will explore the relationship between Common Mongolic, the reconstructed ancestral language of all Mongolic languages on the one hand, and the Qīnghǎi-Gānsù languages on the other. It will investigate how the development of Mongolic lexemes in these peripheral Mongolic languages can contribute to the reconstruction of the earliest forms and later phonetic history of these lexemes themselves, as well as to our knowledge of Common Mongolic phonology as a whole.

Although Qīnghǎi-Gānsù Mongolic has had a long separate development, only incidental word shapes are older than forms documented in Middle Mongol. The evaluation of these languages has widened the base of Common Mongolic, without however increasing its time depth. Most developments of the Qīnghǎi-Gānsù languages can be dated after the Mongol expansion of the 13th century.

Until recently the reconstruction of Common Mongolic lexemes heavily relied on a limited set of languages. In the first place, Written Mongol spellings have long been believed to accurately reflect an older stage of Mongolic. In the second place, the well-known (and politically important) ‘central’ Mongolic languages are generally used: Mongolian proper (including Khalkha and Chakhar), Oirat (including Kalmuck), and Buriat. Middle Mongol sources in several scripts were consulted to add information on specific details, such as the initial **h-* sound, and diphthongs lost in the central languages. Data from non-Mongolic (especially Turkic and Tungusic) languages were often used in support of the resulting reconstructions.

Data from the ‘peripheral’ Mongolic languages, i.e., those that are not in the above-mentioned ‘central’ group, also found their way into comparative Mongolic studies. Poppe (1955) used Dagur, Mongghul, and Moghol.

Materials for Baoan, Dongxiang, and Eastern Yugur were published by Potanin as early as 1893, but the material was quite limited until relatively recently. Poppe (1955) could not yet harness the data from these languages for reconstruction purposes.

Since Poppe’s time we have gained a lot of additional information on these languages, mainly thanks to Chinese and Soviet publications, the delayed results of

¹ The discovery of Kangjia demonstrates that finding new varieties of Mongolic is not out of the question. The term ‘Qīnghǎi-Gānsù (QG) languages’ will only be used here to refer to the peripheral Mongolic languages of these provinces. It thus excludes varieties of central Mongolic (Oirat and Mongol proper) also found here.

the Sino-Soviet expeditions of the 1950s. It was now becoming clear how different the peripheral Mongolic languages are from the central languages, and from one another.

The peripheral languages are not a single subgroup of related languages. There are at least three, but probably four independent groupings: Dagur in the Northeast, Moghol in the Southwest in Afghanistan, maybe all but extinct, and the Shirongol languages in Gānsù and Qīnghǎi provinces. Shirongol is the collective name for the vernaculars gathered under the names Mongghul, Mangghuer (the Monguoric branch), and Baoan, Kangjia, Dongxiang (the Baoanic branch).² Eastern Yugur seems to form a fourth group genetically; similarities between it and the Shirongol languages may be largely due to areal convergence.

Research for this project was started in the hope and expectation that the Qīnghǎi-Gānsù languages would be able to confirm and enhance our knowledge of the ancestral Mongolic language. In general they do provide confirmation of existing reconstruction forms, thus broadening their base.

The value of the Qīnghǎi-Gānsù languages for some phonetic details of Common Mongolic was already known. Both Eastern Yugur and the Shirongol languages preserve a number of old features that are absent from the central languages.

The initial *h*- documented in Middle Mongol survives (in various forms) in all QG languages (as well as Dagur). See 4.13.

The complex vowels **au* and **eü* are partly preserved in both branches of Shirongol (also shared with Dagur). See 3.13.4.

Preconsonantal *-l-* is preserved in Shirongol in a small group of words including **mölsün* ‘ice’ (as in Middle Mongol; this feature was also preserved by Khamnigan and Buriat). See 4.10.3.

Some instances of *q* and *ġ* in Eastern Yugur and the Baoanic languages may support the existence of a vowel harmonic counterpart of **i* (as in Middle Mongol in Arabic script). See 1.5.4.1. and 4.5.2.

Most further lexical and phonetic contributions of the Qīnghǎi-Gānsù (henceforth: QG) languages that will be discussed are due to their relative isolation, which enabled them to preserve features that were lost elsewhere. The QG languages have been separated from the other peripheries, and, more importantly, from central Mongolic. Most QG forms involve regional developments, or the absence of a central Mongolic development.

A related benefit of the QG languages is the lack of influence from Written Mongol, whereas the central languages have been influenced by, and have been influencing, the written forms for centuries. Compare the development of the ordinal suffix **-dUAr* > to *-dUgAAr* or invariable back-vocalic *-dugaar* in the central languages, as opposed to Dagur *-dAAr*, Mongghul *-dar*, Dongxiang *-da*, which developed according to sound laws.

The QG word shapes rarely reveal completely unexpected features of Common Mongolic lexemes. They do provide an additional opinion on the

² The term Shirongol, which encompasses the Qīnghǎi-Gānsù languages except Eastern Yugur, was used by Potanin in the same sense, apart from the fact that Kangjia was not known to him.

reconstruction of words with contradictory forms in the central languages, Middle Mongol and Written Mongol.

The forms found in the QG languages are not necessarily archaic. Individual word shapes deviating from the central languages may preserve archaic features, e.g. *s-* (when corresponding to central Mongolic *š-*), or *d* (central *ǰ* or *ʒ*). In a case like **kabar* ‘nose’ (central **kamar*) it is not obvious which variant is the older.

Due to the word-final accent prevailing in the QG languages they tend to preserve the very vowels that were reduced or elided in the central languages. They thus fulfil a complementary role to central Mongolic in the detection of the presence, location, and quality of vowels.

However, both sets of languages may not only elide unaccented vowels, but also break up original consonant clusters, leading to the appearance of non-etymological vowels. Therefore the Middle Mongol and Written Mongol forms are often indispensable to arrive at the correct reconstruction form. On the other hand, limitations of the writing systems make it unsafe to rely on these old written forms alone.³

In the following cases the QG forms contribute useful evidence, although the added data do not always lead to a definitive reconstruction.

a) the presence of vowels, especially in word-final position, e.g. **kora* ‘poison’ (Written Mongol has a spelling alternation *qoor* ~ *qoora*), **öreele* ‘hobble’ as opposed to Written Mongol *örögel*. However, especially high vowels following *r* tend to be lost, perhaps by absorption into this trill consonant, cf. **jïar(i)* ‘musk’, **saarï* ‘hindquarters’, **siberi* ‘foot sweat’. The QG languages can also be useful for detecting other vowels in non-first syllables.

b) the quality of vowels, more specifically vowel height. In such cases a vowel was certainly present, but its quality could not be established on the basis of the three central standard languages, where the distinction between high and non-high short vowels is largely neutralised in non-first syllables. Like Khamnigan and Ordos, the QG languages show different developments for **sere-* ‘to sense’ and **seri-* ‘to wake up’, **kada-* ‘to sew on’ and **kadu-* ‘to harvest’, **büte-* ‘to be covered’ and **bütü-* ‘to finish’.

In some instances the peripheral languages shed light on ambiguous central Mongolic or Written Mongol forms, or at least shift balance of evidence. In case of the Written Mongol spelling variants *jegüde(n)* ~ *jegüdü(n)* the QG languages support the reality of the latter (in this case contra Ordos *ju:de*). There are many similar cases, in which the peripheral languages favour a certain reconstruction, without establishing that this is the single ancestral form that all modern languages go back to. The reconstruction **kökül* ‘forelock of a horse; braid of hair’ is supported by Ordos *gukul*, but the Eastern Yugur and Mongghul equivalents suggest **kökel*. **olusun* ‘hemp’ is supported by Ordos *ulusu*, but Eastern Yugur and Mongghul suggest **olasun* instead. The reconstruction **eber* or **öber* ‘bosom’ agrees with Eastern Yugur *βer*, Mongghul *yer* and Kangjia *ver* (as opposed to Written Mongol spellings *ebür*, *öbür*). The QG languages also suggest **sini* ‘new’ (spellings *sine* ~ *sini*), again contradicting Ordos *šine*. Ordos *böjōŋ* ‘young hare’ and

³ Shortcomings related to vowel detection include the frequent omission of vowels in sources in Arabic script and the presence of potentially non-existing vowels in sources in Chinese script, as in *zhe-mi-shi* for **jemis* ‘fruit’).

Eastern Yugur *peŋeŋ* suggest **böŋeŋ* with a non-high vowel in the second syllable, while Khalkha and Kalmuck suggest **böŋiŋ*. Such disagreements will be indicated throughout the comparative supplement, see for instance **jebe* (?**jebi*) ‘rust’, **jele* (?**jelü*) ‘rope to which the tethers of cattle are attached’, **kaïnag* ~ **kaïnug* ‘yak’.

1.2. Mongolic Languages

1.2.1. Overview and global classification

Without revisiting the discussion about ‘dialect’ and ‘language’, the following Mongolic linguistic entities can be distinguished. For listings and taxonomies of the central Mongolic dialects I refer to existing publications such as Doerfer (1964) and Janhunen (2003e). Rybatzki (2003) deals with the taxonomy of all Mongolic languages, and summarises previous classifications.

Northeastern periphery

- Dagur

Central languages

- Mongol proper
- Ordos
- Oirat-Kalmuck
- Khamnigan
- Buriat

Western periphery

- Moghol

Southern periphery

1) Eastern Yugur

- Eastern Yugur (Shira Yugur or Nggar)

2) Shirongol

2a) Monguoric

- Mongghul (Huzhu Monguor)
- Mangghuer (Minhe Monguor)

2b) Baoanic

- Baoan (Bonan)
- Kangjia
- Dongxiang (Santa)

This list reflects an approximate genetic subdivision, although it does not do justice to the complexity of the original relationships and subsequent contacts between these languages. Taxonomic remarks in this book will mostly be limited to the internal relationships of the QG languages (see especially 1.5. below).

1.2.2. General characteristics and typology

Several of the assumed features of Common Mongolic are no longer shared by all, or even most, Mongolic languages.

Common Mongolic was an agglutinative SOV language. All affixes were suffixes placed after the stem in a prescribed order. Suffixes typically had a single function, and retained a distinct shape in the chain of suffixes. Most suffixes had several alternants, one of which was chosen based on the phonetic properties of the stem, such as its vocalism (vowel harmonic class) and any final consonant. Common Mongolic vowel harmony involved two classes of vowels. The distinction between the two may have been an opposition between front and back vowels or may have been based on tongue root position. The QG languages do not provide additional evidence to resolve this matter.

Substantives and adjectives were not strictly separated. Pragmatically all non-verbs (except particles) can be considered nouns.

The noun could occur in singular and plural; the latter was indicated by a variety of suffixes. Grammatical and basic spatial relations were expressed by case endings for the genitive, dative, accusative, ablative, instrumental, and comitative. More precise spatial, temporal and logical relations would be expressed by means of postpositions.

Some nouns had two forms, one of which (marked by *-n*) was used attributively, and the other in the predicate, in enumerations, or independently.

The oldest documents show evidence of grammatical gender, as well as some agreement. Modern languages now only preserve masculine and feminine forms for a small number of nouns and adjectives.

Personal and demonstrative pronouns based their case forms on oblique stems rather than the bare stem (which coincides with the nominative). The genitive of the personal pronoun could be placed before or after the noun to act as a possessive marker; the postposed genitives resulted in suffixes only in later stages (and are then placed after the case endings, as in Tungusic, but unlike Turkic). A reflexive possessive referred to the agent of the verb, irrespective of the grammatical person.

The personal pronouns have three persons in singular and plural, without gender distinction; the first person plural distinguished between exclusive and inclusive. The demonstrative pronouns later mostly supplanted the third person personal pronouns.

The verb had several finite tenses. Grammatical person was expressed by the personal pronoun. Only later, and in some languages, these developed into personal endings.

The simple imperative had the same shape as the bare verb stem. Many other imperatives, optatives, etc., existed. There were several participles or verbal nouns. Verb actions were stringed together by means of converbs.

Verbal negation is expressed by means of a large set of negative particles, several of which appear to be based on (defective) negation verbs, again as in Tungusic.

Questions are expressed by means of question particles.

For the inflectional suffixes of the ancestral language I refer to Janhunen 2003b. A selection of derivational suffixes is listed after Chapter 4.

1.2.3. Mongolic lexicon

As many of the typological features mentioned above are shared by other Eurasian languages, the lexicon is the defining trait of the Mongolic languages.

The limited time depth of Common Mongolic means that the reconstruction forms are not very speculative, and each individual item generally resembles the majority of its modern descendants. Differences between the reconstructions of various authors mainly concern conventions and notation.

The young age of the reconstructed language also means that many parts of the vocabulary are known in great detail. We know the specialist terms for male, female, young, infertile horses (of which **agta*, **ajirga*, **baïtasun*, **daagan*, **geün*, **unagan*, **ürie* are listed in the supplement). Another rich category is that of the edible bulbs (including **gogal*, **kaliar*, **kümeli*, **mangir*, **sarimsag*, **songina*, and **sorïsun* listed in the supplement). A large number of body parts is known, including the individual names of most fingers of the human hand (**kuruun* ‘finger’ and **herekei* ‘thumb’ are listed). Some widespread collocations may go back to the Common Mongolic period, such as **kara daru-* ‘to have a nightmare’ (lit. ‘for something black to press’) and **čikin kormai* ‘earlobe’ (lit. ‘ear hem’).

From an etymological point of view several categories can be recognised within the Mongolic lexicon; unique Mongolic items, items with related forms in non-Mongolic languages but with a specific Mongolic shape, and obvious loanwords from non-Mongolic languages. The second category is by some scholars adduced as evidence for a genetic relationship between Mongolic and other language families. This matter will not be touched upon here.

There is no documented stage of Mongolic that precedes the adoption of loanwords from Turkic, Indo-Iranian, and Chinese. Words adopted from these languages have not been systematically excluded from the comparative vocabulary, provided they meet some criteria mentioned below. Thus, the fact that a certain lexeme was listed in the supplement as a ‘Mongolic word’ (old and surviving in several corners of the Mongolic-speaking world) does not exclude the possibility that the same word is also a Turkic word or an Iranian word.

1.3. Source materials

1.3.1. General considerations

The comparative supplement at the end of the book was conceived as a working corpus to investigate to what degree the QG languages contribute to our knowledge of Common Mongolic. In order to be able to appreciate in which cases the QG

languages yielded new insights, equivalents in the other modern Mongolic languages and a selection of Middle Mongol sources were added wherever possible. The selection of lexemes that now constitutes the comparative supplement is meant to amply illustrate the phonetic developments discussed in chapters 3 and 4, provide the sources and page numbers for the words cited in the text, explain the reconstruction forms where necessary, and indicate any regional variants. As a whole it gives an impression of the phonetic diversity and, to some degree, the lexical diversity within Mongolic.

1.3.2. Selection of lexemes

The comparative supplement contains about 1350 items, which is by no means all of the old vocabulary that survives in the QG languages. Some general guidelines have been followed in selecting the listed items.

Nearly all of the included word stems exist in one or more, preferably many, of the Qīnghǎi-Gānsù languages. Within this group, preference has been given to words that are also documented in Middle Mongol and in the central languages.

The list contains examples for the regular (sound law) development of all Common Mongolic phonemes in all positions where they can be found. As some phonemes may behave differently in native and non-native stems (e.g. *š), examples for both groups have been included.⁴ Apart from such regular words, words with an interesting phonetic development or distribution, possibly relevant for classification, have been included as well.

The largest possible number of basic vocabulary items that the reader may look for has been listed, such as personal pronouns, numerals, colour names, animal husbandry terminology, body parts, and kinship terms. Widely occurring words that show little phonetic change (e.g. *kara ‘black’ or *ta ‘you (pl.)’) have also been listed.

Most of the listed items are primary noun and verb stems (the noun category includes numerals, adjectives, adverbs, and pronouns). Derived words have rarely been included, unless the base form was lacking in the relevant languages. The phonetic development of suffixes has mostly been disregarded.

1.3.3. Materials of the QG languages

The focus here is on the contribution that the peripheral languages can make to the reconstruction of Common Mongolic. Languages with a sizeable and searchable corpus were quoted systematically. From the outset it was clear that Eastern Yugur, Monguor, Baoan and Dongxiang should minimally be included. For Monguor both Mongghul (Huzhu Monguor) and Mangghuer (Minhe Monguor) were quoted. For Baoan both the Nantoq subdialect of (Qīnghǎi) Baoan and the Dahejia subdialect of

⁴ While compiling the etymological list I abandoned an earlier idea to exclude all words with known foreign connections including Turkic-Mongolic cognates. This idea was based on the expectation that the QG data would yield a different view of certain aspects of Common Mongolic phonology.

(Gānsù) Baoan were quoted. Kangjia, which is clearly not a variety of Baoan or Dongxiang, was also consulted.

Some other dialects were quoted where they shed light on the history of individual words. This applies to the Ganhetan and Xiazhuang subdialects of Baoan, and the subdialects of the Monguor languages, but also to the hardly known dialects of Eastern Yugur and Dongxiang.

A considerable, and increasing, amount of data is now available for the peripheral languages. This has made it both undesirable and impossible to quote all documented forms here.

A shortcut would have been to take all QG language forms from Sun (1990) and compare those. However, that compilation is not consistent enough to be used uncritically. As a general principle I consulted the most voluminous reliable publications that were available to me. Often these were the volumes from the ‘Hohhot series’ (*Měnggǔ yǔzú yǔyán fāngyán yánjiū cóngshū* = *Series of investigations into the dialects of the Mongolic language family*) and the ‘Peking series’ (*Zhōngguó xiǎoshù mínzú yǔyán jiǎnzhi cóngshū* = *Series of concise grammars of minority languages of China*). Making the data from these sources accessible to a wider audience was also a consideration. Todaeva’s monographs and de Smedt & Mostaert’s Monguor dictionary are present in many libraries, and have been quoted less frequently here.

For most languages one large vocabulary was consulted as the primary source. When a given etymon was not found there, other publications were consulted. Additional forms from other sources are given as well, especially when the phonetic shapes contain additional information about the history of the word. Also words with specific reconstruction problems, or contradictory or hard to interpret forms in other languages, were cited from more publications.

Eastern Yugur forms were cited from Bolčuluu et al. (1984 [1985]) and Junast (1981b), augmented by Bolčuluu & Jalsan (1988) and occasionally by older sources.

Mongghul forms were mainly taken from Khasbaatar et al. (1985 [1986]) and Junast (1981). De Smedt & Mostaert (1933) has been consulted selectively. Lǐ Kèyù’s dictionary (1988) became available to me only later into the project, and I have consulted it sparingly.

Mangghuer materials were taken from various smaller sources: Čenggeltei et al. 1988 [1991], Junast (1981), Junast & Lǐ (1982), Dpal-ldan-bkra-shis, Stuart et al. (1996), as well as the wordlist in Slater (2003).

Dahejia Baoan was primarily cited from Bōkh & Liú (1982).

Ñantoq Baoan was primarily cited from Chén et al. (1985 [1986]).

Other Baoan dialect forms were generally taken from the following sources: Ganhetan from Chén (1995), Jishishan from Li (ms.), Xiazhuang from Bōkh & Chén (1981), incidental forms from other dialects were taken from Chén et al. (1986 [1987]).

Kangjia was cited after Sečenčogt (1999).

Dongxiang was cited after Bōkh et al. 1983. and Liú (1981). The dictionary by Mǎ & Chén (2000) became available to me only recently. It was occasionally cited it when it provided new lexemes or useful variants.

Several of Todaeva's monographs have been used to find further variants: Mongghul and Mangghuer (1973), Baoan (1964), Dongxiang (1961).

Sun's comparative dictionary (1990) has also been selectively used. Potanin (1893) has been used for Eastern Yugur and Shirongol (Baoan, Dongxiang, Wuyangbu (?Weiyuanbu) = Mongghul, and Sanchuan = Mangghuer), especially where he documents words that have since been lost, or provides older phonetic shapes.

1.3.4. Materials of other Mongolic languages

The Written Mongol spellings of each etymon are given wherever they exist. Middle Mongol forms are cited but not necessarily exhaustively. For the modern central Mongolic languages, only the three literary languages Khalkha, Buriat and Kalmuck were consulted systematically. Ordos and Bargu were quoted often. Khamnigan is occasionally quoted when it contributes to the reconstruction of problematic words. Khalkha, Buriat and Kalmuck dialect forms may be included if they were mentioned in the main dictionaries. A systematic evaluation of the many other central Mongolic dialects for historical purposes has yet to be undertaken. Both other peripheral Mongolic languages, Dagur and Moghol, are quoted.

For the old Mongolic documents, a selection of sources in the various scripts has been consulted.⁵

The Secret History and the Huá-Yí Yiyǔ (in Chinese characters with some adaptations) were consulted for all lexemes (after Haenisch 1939, and Mostaert 1977). The Zhìyuán Yiyǔ is incidentally quoted (after Kara 1990).

The Muqaddimat al-Adab (in Arabic script) was consulted for all lexemes (after Haenisch 1939, Mostaert 1977 and Poppe 1938-39). Other sources in Arabic script were quoted selectively, including the 'Leiden Manuscript' and the 'Istanbul manuscript' (after Poppe 1927-28 and Ligeti 1962, 1963), and Ibn Muhanna (Poppe 1937-38). The Rasulid Hexaglot (Golden, et al. 2000) has been quoted, mostly when it contained additional phonetic information.

Forms in 'Phags-pa script are cited when they were listed by Poppe 1957.

For Written Mongol Lessing (1960) was primarily consulted. For the three central standard languages the well-known dictionaries were used: Hangin (1986) for Khalkha, Čeremisov (1973) for Buriat, and Muniev (1977) and Ramstedt (1935) for Kalmuck. Other dictionaries were consulted where necessary.

The primary source for Dagur were Enkhbat's materials, mostly (1984), but also (1983) and (1988).⁶ Additional Dagur forms were mostly taken from Zhòng (1982) and Namcarai & Khaserdeni (1983).

Moghol was cited after Ramstedt (1906) and Weiers (1972), augmented with Moghol words recorded by Ligeti (1963, 1964, 1968, 1974).

⁵ Most of the comparative supplement was prepared in the 1990s. It was not possible to systematically consult newer editions and commentaries on the older stages of Mongolic, such as de Rachewiltz (2004), Saito (2006, 2008), and Tömörtogoo (2006).

⁶ As a native speaker Enkhbat was able to better distinguish real Dagur forms from central Mongolic forms and Mongolic words reborrowed from Manchu and northern Tungusic.

1.3.5. Reconstruction of individual lexemes

The reconstruction forms are listed alphabetically in the comparative supplement. Readers familiar with Written Mongol will have little trouble finding a given word, because in many respects spellings in Uigur script resemble Common Mongolic quite closely.

The reconstruction of individual words is often straightforward because of the limited time depth of Common Mongolic. However, the phonetic developments in the various languages have occasionally altered related etyma beyond (immediate) recognition.

Reconstruction problems mostly involve contradictory information from the various languages or subgroups. The original word shapes need to be distinguished from secondary developments. Particular attention should be paid to the following:

- a) Establishing the location and quality of the vowels in the stem (vowel detection).
- b) Establishing whether consonant strength is primary or secondary. This includes distinguishing primary **h-* from secondary *h-* (and allied consonants such as *f-*).
- c) Establishing whether first members of consonant clusters are original or later additions.

Morphological dissection occasionally helps to connect related words, e.g. **haakai* and **haaljin* ‘spider’ reveal a root which is not known to occur independently.

1.3.6. Semantics

The meanings of the listed lexemes in individual languages are generally omitted when they agree with the meaning or set of meanings assumed for Common Mongolic. Meanings may be given if a semantic difference between the languages casts doubt on the identity of the forms. Semantic developments may also be mentioned and discussed when they are typical of a certain area or subgroup.

1.4. Phones of the cited languages

1.4.1. General considerations

This section contains the vowel and consonant inventories of the peripheral Mongolic languages. By and large these surveys will be organised as phonologies, but in the interest of historical transparency and ease of comparison of the data, I will on some occasions deviate from previous analyses.

The data on which I based my reconstructions and analyses are included in the comparative supplement, all provided with page numbers of the original publications. The published notation of each item can thus be retrieved.

For most Mongolic languages spoken in Gānsù and Qīnghǎi provinces there are now at least five phonological analyses which can be consulted in preparing this survey: those found in Todaeva’s monographs; the ‘Hohhot series’; the ‘Peking

series'; various authors in *The Mongolic languages*; and Svantesson et al. in *The phonology of Mongolian*.⁷

To enable convenient comparison of the data, some degree of standardisation of the notation was necessary. In doing this, a balance had to be struck between phonetic reality, phonological analysis and readability, and between diachronic transparency and synchronic comparability.

Given the goals of this book, diachronic transparency was generally given preference over phonological 'tidiness'. Marginal phonemes and speech sounds whose distribution is synchronically and/or diachronically unclear have not been 'sanitised', and have mostly been left intact (examples include the vowels *i* and *ə* in Eastern Yugur, the vowel *ɛ* in Kangjia, the retroflex versus the alveopalatal consonants in Mangghuer, *g* versus *ǰ* in Dongxiang, unexplained vowel lengths). Such occurrences may be explained in the future, and may even reveal details of historical development.

Unfortunately it was not feasible to acknowledge and endorse the emerging Pinyin-based writing systems of Mangghuer, Mongghul and Dongxiang. The reader would have had trouble generating phonetic realisations from the orthography⁸, e.g. Dongxiang <*xien*> is pronounced [*ʃien*] with *-n*, but <*khan*> is pronounced [*qʰaŋ*] with *-ŋ*. There are also discrepancies between the orthographies, e.g. the syllable <*zho*> is to be read [*ʃuo*] in Dongxiang, but Mangghuer spelling does not allow this non-Pinyin syllable, and prescribes the notation <*zhuo*> whether or not a diphthong is pronounced.

Following a convention in Chinese publications about the QG languages, weak consonants have been written with graphemes that are traditionally associated with voiced consonants, strong consonants with graphemes associated with voiceless consonants, e.g. *d* stands for [*t ~ ʈ*], and *t* for [*tʰ*]. However, affricates have been written here with single characters rather than ligatures, e.g. *č* instead of *tʃ*.

Phones restricted to borrowed lexemes are not included below. Some of the marginal phonemes occurring in native lexemes, and phones whose phonemic status is unclear, have been mentioned.

1.4.2. Non-QG languages

Apart from some superficial changes in the notation (such as the replacement of Haenisch's *ś* and *h* by *š* and *q*) the different spellings for Sino-Mongolian have been left as deciphered by the various authors. Some of the forms in documents in Arabic script have been read differently here, but such cases have been indicated.

Khalkha, Buriat, and Kalmuck have been transliterated from the Cyrillic spelling, with only incidental further clarifications (e.g. whether <*ě*> should be read *yo* or *yö*). Ramstedt's Kalmuck spellings have been slightly altered. Dagur has been cited in a form close to the (phonetic) notation of Enkhbat. Moghol has been cited from the various sources without attempting to standardise the notation.

⁷ For a survey of earlier phonological analyses of the QG languages see Svantesson et al. (2005: 149-154), and the various chapters in Janhunen (2003).

⁸ The Mangghuer orthography is probably easiest to read, at least for those acquainted with Chinese written in Pinyin.

1.4.3. Eastern Yugur phones

The number of Eastern Yugur native vowel phonemes is hard to determine. Bolčuluu & Jalsan (1990) distinguished ten vowels /a/, /e/, /ə/, /i/, /ɔ/, /o/, /ö/, /u/, /ʉ/, /ü/, all with long counterparts. The independence of some of these phones cannot be demonstrated. On the other hand the analyses of Tenišev & Todaeva and Junast seem to be over-simplified. There are probably seven vowel phonemes /a/, /e/, /ə ~ i/, /ɔ/, /ö/, /u/, /ü/, with long counterparts. Short *i* is a relatively rare allophone of /ə/. Both synchronically and diachronically it seems likely that [o] is a defronted allophone of /ö/ and that [ʉ] is a centralised allophone of /ü/.

Eastern Yugur consonants in native lexemes:

	labial	apical	palatal	velar	uvular	phar./lar.
strong plos.	<i>p</i>	<i>t</i>		<i>k</i>	<i>q</i>	
weak plos.	<i>b</i>	<i>d</i>		<i>g</i>	<i>ġ</i>	
strong affr.		<i>c</i>	<i>č</i>			
weak affr.		<i>ɟ</i>	<i>ǰ</i>			
strong fric.		<i>s</i>	<i>š</i>		<i>χ</i>	<i>h</i>
weak fric.	<i>β</i>			<i>ɣ</i>	<i>ġ</i>	
nasal	<i>m</i>	<i>n</i>		<i>ŋ</i>		
lateral		<i>l</i>				
trills etc		<i>r</i>				
approx.			<i>y</i>			

The plosives *g* and *ġ* and the fricatives *ɣ* and *ġ* are only carefully distinguished by Junast; Bolčuluu, while acknowledging the existence of the fricatives, only uses the plosives in his notation.

The voiceless consonants *ʃ*, *ɲ*, and *ɣ* can be phonologically analysed as sequences /hl/, /hn/ and /hy/, which mostly agrees with their diachronic background.

1.4.4. Mongghul phones

Mongghul has a compact vowel system of five short vowels /a/, /e/, /i/, /o/, /u/, with long counterparts. The phoneme /i/ mostly appears as a centralised [ə].

Mongghul consonants in native lexemes:

	labial	apical	alveopal.	retrofl.	velar	uvular
strong plos.	<i>p</i>	<i>t</i>			<i>k</i>	
weak plos.	<i>b</i>	<i>d</i>			<i>g</i>	<i>ġ</i>
strong affr.		<i>[c]</i>	<i>č</i>	<i>[ɕ]</i>		
weak affr.		<i>ɟ</i>	<i>ǰ</i>	<i>[ʑ]</i>		

Mongghul consonants (continued)

strong fric.	<i>f</i>	<i>s</i>	<i>ś</i>	<i>ʂ</i>	<i>x</i>
nasal	<i>m</i>	<i>n</i>			<i>ŋ</i>
lateral		<i>l</i>			
trills etc		<i>r</i>			
approx.	<i>w</i>		<i>y</i>		

1.4.5. Mangghuer phones

The Mangghuer vowel system has five short vowels, /a/, /e/, /i/, /o/, and /u/, perhaps six if one counts /ə/ as a separate phoneme. Incidental vowel lengths survive.

Mangghuer consonants in native lexemes:

	labial	apical	alveopal.	retrofl.	velar	uvular
strong plos.	<i>p</i>	<i>t</i>			<i>k</i>	
weak plos.	<i>b</i>	<i>d</i>			<i>g</i>	<i>g̃</i>
strong affr.		<i>c</i>	<i>č</i>	<i>č̣</i>		
weak affr.		<i>ʒ</i>	<i>ž</i>	<i>ẓ̌</i>		
strong fric.	<i>f</i>	<i>s</i>	<i>ś</i>	<i>ʂ</i>	<i>x</i>	
nasal	<i>m</i>	<i>n</i>			<i>ŋ</i>	
lateral		<i>l</i>				
trills etc		<i>r</i>				
approx.	<i>w</i>		<i>y</i>			

r is often pronounced [ʐ] as in Mandarin, but will not be analysed as a fricative here.

1.4.6. Dahejia Baoan phones

Dahejia Baoan has six vowels: /a/, /e/, /ə/, /i/, /o/, /u/. Dahejia Baoan consonants in native lexemes:

	labial	apical	alveopal.	retrofl.	velar	uvular	phar./lar.
strong plos.	<i>p</i>	<i>t</i>			<i>k</i>		
weak plos.	<i>b</i>	<i>d</i>			<i>g</i>	<i>g̃</i>	
strong affr.			<i>č</i>				
weak affr.		<i>ʒ</i>	<i>ž</i>				
strong fric.	<i>f</i>	<i>s</i>	<i>ś</i>	[ʂ]		<i>χ</i>	<i>h</i>
nasal	<i>m</i>	<i>n</i>			<i>ŋ</i>		
lateral		<i>l</i>					
trills etc		<i>r</i>					
approx.	<i>w</i>		<i>y</i>				

The fricative $\mathfrak{s}$ occurs instead of $\mathfrak{z}$ in related (Gansu) Baoan dialects.

1.4.7. Ñantoq Baoan phones

Ñantoq Baoan has six vowels $/a/$, $/e/$, $/ə/$, $/i/$, $/o/$, $/u/$, with long counterparts for all except $/ə/$.

Ñantoq Baoan consonants in native lexemes:

	labial	apical	alveopal.	retrofl.	velar	uvular	phar./lar.
strong plos.	p	t			k		
weak plos.	b	d			g	$\mathfrak{g}$	
strong affr.		c	$\mathfrak{c}$				
weak affr.		$\mathfrak{z}$	$\mathfrak{z}$				
strong fric.	f	s	$\mathfrak{x}$	$\mathfrak{s}$		$\mathfrak{x}$	h
weak fric.			$\mathfrak{z}$				
nasal	m	n			η		
lateral		l					
trills etc		r					
approx.	w		y				

In native words $\mathfrak{s}$ occurs as the first member of clusters, while $\mathfrak{x}$ (in fact a voiceless palatal fricative), is found before vowels. In Mongolic words, $\mathfrak{z}$ is occasionally found in medial position. Svantesson is correct in viewing $\mathfrak{x}$ and $\mathfrak{s}$ as allophones in native words.

1.4.8. Kangjia phones

Kangjia is described as having nine vowels $/a/$, $/e/$, $/ə/$, $/i/$, $/i:/$, $/ɔ/$, $/o/$, $/u/$, $/u:/$, but the status of $/ə/$ and $/i:/$ is not entirely clear. Uniquely among Shirongol languages, Kangjia largely preserves the distinction between back $*o$ and $*u$ and front $*ö$ and $*ü$, thus making it impossible to posit the merger of the two sets for Proto Shirongol, or even proto Baoanic. Some long vowels remain, but do not constitute a system of oppositions.

Kangjia consonants in native lexemes:

	labial	apical	palatal	velar	uvular	phar./lar.
strong plos.	p	t		k	q	
weak plos.	b	d		g	$\mathfrak{g}$	
strong affr.		c	$\mathfrak{c}$			
weak affr.		$\mathfrak{z}$	$\mathfrak{z}$			
strong fric.	f	s	$\mathfrak{s}$		$\mathfrak{x}$	h
weak fric.	v	z		y	$\mathfrak{g}$	
nasal	m	n		η		

Kangjia consonants (continued)

lateral	<i>l</i>
trills etc	<i>r</i>
approx.	<i>y</i>

The status of *ɭ* is synchronically and diachronically unclear.

1.4.9. Dongxiang phones

The Dongxiang vowels are /a/, /e ~ ə/, /i/, /ĩ/, /o/, /u/. Depending on the description, *ĩ* is found after apicals *s*, *c*, *ʒ* and retroflexes *ɕ* *ʑ* *ʐ* (= medial *r*), and/or in the vicinity of uvulars.⁹

Dongxiang consonants in native lexemes:

	labial	apical	alveopal.	retrofl.	velar	uvular	phar./lar.
strong plos.	<i>p</i>	<i>t</i>			<i>k</i>	<i>q</i>	
weak plos.	<i>b</i>	<i>d</i>			<i>g</i>	<i>ġ</i>	
strong affr.		<i>c</i>	<i>ɕ</i>	<i>ɕ</i>			
weak affr.		<i>ʒ</i>	<i>ʑ</i>	<i>ʑ</i>			
strong fric.		<i>s</i>	<i>ʃ</i>	<i>ʃ</i>	<i>x</i>		<i>h</i>
weak fric.				<i>[ʐ]</i>		<i>ġ</i>	
nasal	<i>m</i>	<i>n</i>			<i>ŋ</i>		
lateral		<i>l</i>					
trill		<i>r</i>					
approx.	<i>w</i>		<i>y</i>				

Medial *-r-* is pronounced *ʐ* in some vowel environments, notably following /i/.¹⁰ Final *-r* (pronounced *ɹ*) is marginal in Suonanba, and is often analysed as a part of the preceding vowel (Bökh adds the ‘rhotic vowel’ *ə̃* to his system). In other Dongxiang dialects syllable-final *r* occurs frequently. Liú (1981:13) has a velar fricative *ɣ* rather than a uvular one.

⁹ It may also be found in other environments, as in Dgx (according to Liú’s analysis) *tīǰi* < **toǰi* ‘button’, and *dalī* < **dalū* ‘shoulder’. The other sources record these words differently.

¹⁰ The sequence *-ri* in Dgx *qiri*- < **gar*- ‘to exit’, *zawari*- < **jalbari*- ‘to beg’ does not have a retroflex pronunciation. Ma & Chen write these words with <ei>, i.e., *-ai*. As such a pronunciation is not confirmed by other descriptions, this may be a spelling convention to indicate that *-r-* does not have a retroflex pronunciation.

1.5. Relationships of the Qīnghǎi-Gānsù Languages

The driving forces behind the development of the QG languages have been a large degree of isolation from the main body of Mongolic, and increasingly intensive contact with neighbouring non-Mongolic languages. The isolation and contact may have come about more or less simultaneously. Therefore the main difficulty in determining the exact relationships between the QG Mongolic languages is the fact that they are not only related, but also belong to the same *Sprachbund*. The problem is not so much a scarcity of similarities, but the fact that it is not always possible to distinguish between inherited similarities and those acquired by contact. Sound laws and other systematic changes that can be used for classification purposes are few. Moreover, there is not a single Qīnghǎi-Gānsù *Sprachbund*. There are several sets of isogloss bundles, each encompassing a different set of languages. Each bundle of isoglosses with a similar outline can be viewed as a mini-*Sprachbund*. Obviously a language can belong to several mini-*Sprachbünde*. There is an area encompassing Eastern Yugur, all of Shirongol Mongolic, as well as the Turkic languages Salar and Western Yugur, certain Amdo dialects and Northwest Mandarin dialects. There are also smaller ones. Eastern and Western Yugur form a mini *Sprachbund*. Baoanic and Salar form another one. There is a larger Amdo Tibetan-Shirongol *Sprachbund* excluding Dongxiang, and possibly a Baoanic-Mangghuer one excluding Mongghul.

As a consequence it seems to be impossible to determine whether the modern QG languages were a separate branch of Mongolic before settling in this area. Likewise it cannot be established whether the Monguoric and Baoanic branches of Shirongol entered the QG area as already distinct entities.

In the following pages I will concentrate on the internal division of the QG languages. Shared innovations are known to be the best classification arguments, and within that category shared *irregular* innovations seem to be most useful. Most of the arguments used involve phonetic shifts, morphological differences such as stems with different endings, and lexical differences including semantic and functional changes of existing lexemes. Purely morphological and syntactical arguments will only incidentally be referred to. See Rybatzki (2003b), whose list includes items of all types.¹¹

Genetic classification features should be non-trivial, not easily repeatable by several languages independently, and not easily reversible. They should not be obviously adopted from neighbouring languages. They are preferably not a by-product of another similarity (e.g. the reduction of unaccented vowels and the appearance of initial consonant clusters, both of which are due to the final accent). Such features are manifestations of, or evidence for, a certain development, but should not be counted as additional similarities.

Sound laws and other systematic changes are not necessarily informative in the context of classification. I will attempt to determine for each feature that is shared by two or more languages, and may therefore be of classificatory value, whether it was inherited from a shared ancestor, or developed separately under the influence of a neighbouring language.

¹¹ Not all of the phonetic criteria listed by Rybatzki will be repeated here, which is not a statement about their validity.

1.5.1. Features uniting all of QG Mongolic

The features shared by Eastern Yugur and the Shirongol languages are mostly of a very general nature, and most are not exclusive to the QG region.

Shared phonetic developments include: reductions following from the final accent, a tendency to simplify or abolish vowel harmony, weakening and strengthening of consonants depending on the phonetic environment, and, related to this, vowel devoicing phenomena, the development of secondary *h*- and a new phoneme *p*-.

Most classification arguments involve the irregular development of individual word shapes.

1.5.1.1. Final accent and consequences

Word-final accent is shared by all QG languages. This feature is also found in Moghol. While this clearly sets these languages apart from central Mongolic and Dagur, it is not clear which group of languages is the innovative one. It is possible to assume that the accent originally fell on the first vowel, and to ascribe the final accent in QG and Moghol to the influence of neighbouring Turkic languages, but in the case of the Shirongol languages it is not obvious that Turkic influence was sufficient for this feature to be adopted.

There are exceptions in the QG languages. In Eastern Yugur some word structures trigger changes in the position of the accent. In Baoan most exceptions involve compounds, which apparently carried the main accent on the last syllable of the first member, e.g. Baoñ *'nuda* 'today' < **ene üdür* 'this day', *ta 'waraŋ* '50' < **tabun harban* 'five tens', *'debgə-* 'to blink' < Amdo *hdeb* + **ki-*. In Dongxiang exceptions may occur in words whose last vowel has ended up in the penultimate syllable due to epenthesis, e.g. *'basi* 'tiger' < **bars*, *'bosi* 'fabric' < **böš*.¹² Loanwords from Arabic may have non-final accent, e.g. *'saɣari* 'early morning meal' from Arabic *saḥar*, *'aɣili* 'intellect' from Arabic *'aql*.

Exceptions are also seen in Chinese loanwords, where non-final accents are perhaps the result of attempts to deal with tonal differences, and prevent homophony, e.g. Dgx *ʃi'zi* < *shīzi* 'lion' as opposed to *'ʃiʒi* < *shīzi* 'persimmon'.¹³

A corollary feature of the final accent are reduction and loss of unaccented vowels. This mostly affects the vowels of the first syllable, and, in case of longer stems, the vowels of medial syllables. This development is also observed in Dongxiang, the most conservative QG language in this regard. The following examples illustrate how several QG languages may have similar tendencies but with different outcomes. Different vowels are elided in various lexemes and dialects due to reasons that are not fully understood, but possibly involving accent or differences in vowel

¹² This may be more common than is indicated in the sources. In several words, such as *bosi* < **boas* 'pregnant', initial stress would be expected but is not indicated. In verbs, connective vowels that are reanalysed as the final vowel of the stem do take the accent.

¹³ These examples are from Bökh's dictionary. See Liu (1981:17) for more examples. The opposition between 'lion' and 'persimmon' has been resolved in the same way in Ganhetan Baoan. For this and other examples see Chen (1995:131-132).

quantity (phonemic or otherwise). In all languages, elision is restricted by constraints on syllable structure, so that it cannot lead to inadmissible consonant sequences.

EYu	Dgx	CM	
<i>aiğa</i>	<i>iğa</i>	* <i>ayaga</i>	bowl
<i>ömlö</i>	<i>məlie</i>	* <i>emüne</i>	front
<i>ondor</i>	<i>niudu</i>	* <i>ene üdür</i>	today
<i>nanda ~ danda</i>	<i>namadə ~ madə</i>	* <i>namada/*namadu</i>	to me
<i>dasära-</i>	<i>cıra- ~ sıra-</i>	* <i>tasura-</i>	to break (intr)
<i>adla-</i>	<i>adula-</i>	* <i>aduula-</i>	to tend animals
<i>halağan</i>	<i>haŋga</i>	* <i>halagan</i>	palm

These discrepancies show that individual cases of vowel loss are not reconstructable for older stages, although the general tendency may have developed quite early.

In most languages vowel elision results in consonant sequences. Most notably all QG languages developed initial consonant clusters in native words, something completely unknown in remaining Mongolic, but common in Amdo Tibetan. In Dongxiang, the only QG language that did not undergo Tibetan influence, this phenomenon is only seen in a small number of words that can also be analysed differently (by means of devoiced vowels).

1.5.1.2. Simplification of the vowel system

Only few developments (innovations) are shared by all of QG. The tendency for **i* to be retracted and lowered to *ə* is seen in all QG Mongolic languages, as well as in neighbouring Turkic languages and Amdo Tibetan. The reduction of the number of rounded vowels cannot be assumed for Proto Shirongol, as Kangjia preserves all four of them.

Vowel devoicing phenomena are connected to strong consonants.

Distinctive vowel quantity was largely lost in Mangghuer, Kangjia and Dongxiang, and in some varieties of Baoan, but even there a number of words with historically correct relictual vowel lengths tends to be preserved. Monguoric and Baoanic are thus both descended from ancestors with phonemic vowel length.

In a very general sense the QG languages share the reduction of the vowel harmony system. However, the changes in individual languages differ in the details.

In Eastern Yugur, vowel harmony in stems may be disturbed by secondary developments, but vowel harmony in non-high vowel suffixes (with **A*, in Eastern Yugur represented by the alternation *a/e/ɔ/ö*) remains largely productive. The vowels of the original high vowel suffixes (with **I* and **U*) tend to merge into a neutral *ə* (which phonologically remains a high vowel).

In Monguoric vowel harmony has broken down both in stems and in suffixation. The front rounded vowels **ö* and **ü* merged with their back counterparts **o* and **u*. Nevertheless, the former harmonic constraints are clearly visible in many existing primary and derived stems, as in the Mongghul derivations with the common verbalizer *-la* such as *gar.la-* ‘to grab with the hand’, *ide:.le-* ‘to develop pus’, *noko.lo-* ‘to make a hole’, or the causative *-GA* in *kur.ge:-* ‘to deliver’,

sur.ɡa:- ‘to teach’. Suffixes were later generalised by picking one of the original alternants, usually that with *a*.¹⁴

In Baoan and Dongxiang there are also only two rounded vowels left, but as Kangjia preserves four, generally corresponding to the four rounded vowels of CM, Proto Baoanic must have featured the larger system. Former constraints survive to some extent into modern Baoan and Dongxiang, e.g. the denominal verb suffix **-ra* still occurs in three harmonic variants in existing derivations, as in Dgx *xula.ra-* ‘to become red’, *kugie.re-* ‘to become blue’, *noġo.ro-* ‘to become green’. In modern Dongxiang the suffix as such is still productive (mainly seen on stems adopted from Chinese), but it has now taken on the invariable shape *-ra*. The verbalizers *-dA* and *-lA* also occur in three variants in existing formations, and *-lA* is even found in two harmonic variants in modern derivations from Chinese words (see Liu 1981:23).

1.5.1.3. Phenomena related to consonant strength

In a general sense consonant strengthening and weakening phenomena are shared by the Mongolic languages as well as the Turkic languages of the area, Salar and Western Yugur. This involves the strengthening (aspiration) of weak consonants or weakening (aspiration loss) of strong consonants, mostly in word-initial position under the influence of the consonant starting the second syllable. In QG Mongolic, this type of distant assimilation only seems to take place when the first syllable has a simple vowel. Most languages also feature weakening of medial consonants, which is an unconditioned tendency.

Shifts in consonant strength are one of the shared QG tendencies that cannot be ascribed to Chinese or Tibetan influence, as in these languages the distinction between aspirated and unaspirated consonants carries great phonological weight.¹⁵ Although Mongolic, unlike Turkic, has an opposition between strong **t*, **č*, **k* and weak **d*, **j*, **g* in initial position, changes from strong to weak or vice versa have rarely resulted in a loss of distinctiveness of individual stems, perhaps because Mongolic stems are usually bi- or trisyllabic. The effects include the following:

Assimilatory initial strengthening, as in EYu *qagča* < **gagča* ‘alone’.

Assimilatory strengthening followed by medial weakening. Examples of this can be found in all of QG, but it is most systematically applied by Mongghul, as in *təʒin* < **döčün* ‘forty’.¹⁶

Dissimilatory initial weakening, as in BaoŃ *ġəśəŋ* < **kosiuŋ* ‘beak’.¹⁷

Assimilatory initial weakening, a shared derived feature of Eastern Yugur and Shirongol. For unknown reasons it is largely limited to the weakening of initial **k-*

¹⁴ The frequently used Mongghul perfect converb *-AA(nu)* exceptionally preserves productive alternants with *a*: ~ *e*: ~ *o*:

¹⁵ In Tibetan this is also important morphologically, as aspiration plays a role in distinguishing verb tenses.

¹⁶ The combined effect of this sequence of events looks like a kind of ‘metathesis of aspiration’, hence Svantesson’s term ‘flip-flop’.

¹⁷ This type of conditioned initial weakening is known from central Mongolic languages including Chakhar and Ordos.

when the second syllable starts with **d* or **ʃ*, as in **kada* > *gada* ‘rock’. This does not have a parallel in Turkic and Amdo Tibetan.

Assimilatory medial strengthening, the strengthening of a medial consonant by the strong initial, may explain the Nantoq Baoan forms *ćixaŋ* < **čagaan* ‘white’, *kunjaŋ* < **köŋgen* ‘light’, *kuntə* < **kündü* ‘heavy’.

Strengthening has given rise to two further phenomena shared by the QG languages. Firstly, the weak consonant **b-* has developed a strong counterpart *p-* [*p^h*] that did not exist in the Common Mongolic phoneme inventory. It is also caused by a medial strong consonant, as in Dgx *puʒalu-* < **bučal-* ‘to boil’. The strengthening of **b* to *p-* is also found in neighbouring Western Yugur and Salar.

Secondly, an initial *h-* developed in many words originally starting with vocalic onset. This secondary *h-* has developed the same modern pronunciations as primary CM **h-*, and the two are synchronically indistinguishable. It arose under the influence of the strong consonant starting the second syllable, leading to the devoicing of vocalic onset + vowel, as in MgrM *fʒu* ~ *sʒu* < **hʏsun* < **usun* ‘water’.

Medial preaspiration of consonants is documented for Eastern Yugur and Baoanic, but only in words with original intervocalic **-k-*, as in EYu *qayqai*, Dgx *qixgəi* (see 4.5.3.).

Further consonant developments are discussed in 4.1.3. and the individual consonant sections in chapter 4.

1.5.1.4. Lexical features

The peripheral languages retain a number of lexemes from the Middle Mongol period that were lost in the central languages, creating the (false) impression of a division between Middle Mongol and the peripheral languages on the one hand, and central Mongolic on the other. Examples include **a* ‘they’ (not listed), **biljür* ‘small bird’, **činaida*, **činar*, **činajī* ‘the day after tomorrow’, **daŋgal* ‘lump, clod’, **düre-* ‘to sell’, **haul-* ‘to run’, **heil-* ‘to leave’, **hüü-* ‘to stink, rot’, **i* ‘s/he, it’, **jilsun* ‘glue’, **koa-* ‘to wither’, **najir* ‘summer’, **ödme* ‘bread’, **siür-* ‘to sweep’, **temgü-* ‘to pick up’, **urijün urijī üdür* ‘three days ago’, **ünügün* ‘kid (goat)’, **yama* ‘thing’. These items are more properly viewed as normal words, as several of them were only lost once, in the ancestor of the central Mongolic languages. In many instances related forms from the same roots actually exist in the central languages.

The QG languages also feature regional words that are not attested in Middle Mongol sources. As they do tend to have equivalents in neighbouring Turkic languages, they may ultimately be of non-Mongolic origin. Regional items found in Shirongol include **hayag* ‘handful’, **hurbī* (?) ‘sling’, and **ütergen* ‘threshing floor’. Eastern Yugur and Monguoric share **kirbei* ‘edge’. A regional word found in Eastern Yugur, both Monguor languages, and Dongxiang, is *mila* ~ *mula* ‘small’.

Further regionalisms will be discussed in the following pages. I will focus on words that do exist in central Mongolic and/or the other peripheries, but whose forms in the QG languages suggest a shared development of classificatory importance. The QG languages are sometimes in agreement with Middle Mongol, suggesting that they preserve the original form, while in other cases they are innovative. It is not always possible to determine which of the recorded variants is

the older one. Usually it may simply have been a matter of choosing from among a pool of pre-existing variants. Examples include the following (the actual distribution is more complicated, see the comparative supplement):

QG suggests	MMo suggests	central suggests	
<i>*buli-</i>	<i>*buli-</i>	<i>*bulia-</i>	to snatch
<i>*kabar</i>	<i>*kabar</i>	<i>*kamar</i>	nose
<i>*kadaar</i>	<i>*kadaar</i>	<i>*kajaar</i>	bit
<i>*kürjeg</i>	<i>*kürjeg</i>	<i>*kürje</i>	spade
<i>*tüükü</i>	<i>*tüükü(i)</i>	<i>*tüükei</i>	raw

Examples of this type are few, and as the isoglosses for the various items do not overlap, each case may suggest a different subdivision of the language family. In some cases, such as **kadaar*/**kajaar* ‘bit’, the former seems to be a ‘peripheral’ retention, and the latter a central innovation.

In others, such as **kabar*/**kamar* ‘nose’ and **buli-*/**bulia-* ‘to snatch’, the former is shared by Moghol and QG, the latter by the central languages and Dagur, suggesting a North-South division. This seems to be supported by other features, such as initial accent in the North and final accent in the South, and preservation of the distinction between attributive and enumerative numerals in the North, lost in the South. However, the impression of a North-South division is not confirmed by Middle Mongol, which may side with either region, or may be indecisive (e.g. Middle Mongol accent is not known).

Occasionally, two variants of the same etymon were already recorded in Middle Mongol sources, such as **könerge* ~ **körenge* ‘yeast’, without a predictable distribution. In this case the QG languages preserve the former variant, recorded in ‘Phags-pa. However, with regard to metathetic alternation, the QG languages do not always point in the same direction, see **magalai* ~ **malagai* ‘hat’.

In case of **yaan* and **yaun* ‘what’, QG favours the former variant, recorded in the Muqaddimat al-Adab. The same applies to the derived forms **yama* and **yauma* ‘what’.

In case of **čila-* ~ **čile-* ‘to be tired, to tingle’, the former variant, found in QG, is not attested in Middle Mongol, although one expects it to be the older variant (the reverse shift of harmonic class is unusual).

The verb ‘to laugh’ has the shape **hinie-* both in the QG languages and Dagur, suggesting that it may predate the form **inie-* supported by Middle Mongol.

In a couple of words, the QG languages share an (irregular) phonetic innovation, including the following:

QG suggests	CM as otherwise known	
<i>*doli-</i>	<i>*dolia-</i>	to lick
<i>*go:r</i> < <i>*koar</i>	<i>*koyar</i>	two
<i>*no:r</i>	<i>*noir</i>	sleep
<i>*örle</i>	<i>*öreele</i>	hobble; half
<i>*emele</i>	<i>*emüne</i>	front

Less dramatic potential classification arguments are provided by instances of unexpected presence (preservation or addition) of *-n*, as in **teerme(n)* ‘mill’, or of its unexpected absence, as in **morin* ‘horse’. Monguoric, which tends to drop the *-n* of substantive nouns in the nominative, is less informative in this regard.

1.5.1.5. Semantic and functional features

Semantic and functional developments also yield a number of classification criteria.

A semantic difference involves **kainag*, which denotes the yak in the QG languages, but a yak hybrid in the central languages. In QG the verb **kaïla-* means ‘(for animals) to make a noise’, but elsewhere it mostly means ‘to shout, to cry’. The verb **sïmtara-* means ‘to thaw, melt’ in Shirongol, but ‘to become blunt, calm, weak’ in central Mongolic. CM **keüken* ‘child’ now specifically denotes female children in central Mongolic, but male children in the QG languages.

A functional difference has developed with regard to **ïrgen* ‘people’. It has become a third person personal pronoun in Eastern Yugur and all of Shirongol. **ejen* ‘master’ has become a reflexive pronoun in Eastern Yugur and Monguoric, and a third person pronoun in Baoan. Several Shirongol languages have developed third person pronouns from **nögee* ‘other’ and first or third person pronouns from **öer-* ‘self’ (see 1.5.4.2.).

The numeral **nige* ‘one’ has come to function as an indefinite article and/or singular marker in Monguoric, Baoan and Kangjia.

1.5.2. Peripheral opinions on known Central Mongolic variants

The central languages display a number of correspondences that do not fall under the usual soundlaws. The distribution of variants among the three literary central languages is not always the same, and often goes against the traditional division between Eastern and Western Mongolic.

The fact that there are Khalkha-Buriat, Buriat-Kalmuck, and Khalkha-Kalmuck parallels may point that (at least) a three-way split must be assumed leading to the present three dialect groups. Unfortunately it is generally impossible to determine which variant is the innovative one.

The QG languages in some cases agree with Khalkha, in others with Buriat and Kalmuck, in yet others only with either Kalmuck or Buriat. All of this suggests that several variants coexisted in CM, and that each subgroup came to favour one of them. Distribution of the variants in modern languages certainly does not suggest a binary split of CM into Eastern and Western subgroups.

The preconsonantal *-l-* in the group **čaalsun* ‘paper’, **jïlsun* ‘glue’, **mölsün* ‘ice’, and **sölsün* ‘gall bladder’, was preserved in Shirongol, as in Buriat and Khamnigan, but lost in Eastern Yugur, as in Khalkha and Kalmuck.

Both Eastern Yugur and Shirongol agree with Khalkha on the group **eriün* ‘chin’, **kuruun* ‘finger’, **niruun* ‘back’, rather than the Kalmuck and Buriat forms

with additional *-g-*, suggesting the forms **erigün*, **kurugun*, **nirugun* for these words.¹⁸

In case of the morphological variants **gutusun* and **gutul* ‘boots’, the former form, typical of Kalmuck, is found both in Eastern Yugur and Shirongol, Kangjia excepted. In case of **hargal* and **hargasun* ‘dung’ all QG languages have the ‘Eastern’ variant ending in *-l*.

In case of **sü(n)* ~ **üsün* ‘milk’, Eastern Yugur shares the latter variant with Kalmuck and Ordos, while the word is lacking in Baoanic, and indecisive in Mongghul.

In the case of **turuun* and **tuura(i)* ‘hoof’, the former is found in both Dagur and Eastern Yugur, as in Buriat and Kalmuck; the latter in Khalkha and Ordos.

The verb **hogtal-* ‘to fell’ has preserved its preconsonantal *-g-* in Eastern Yugur as it has in Khalkha, whereas Buriat and Kalmuck reflect a form **otal-* which lost the *-g-*. The Shirongol forms are uninformative as their first syllables are too abraded.¹⁹

All QG languages seem to support the form **öčügedür* (< **öčigen üdür*) ‘yesterday’, also found in Khalkha and Ordos, whereas Buriat, Khamnigan and Kalmuck suggest **öčügeldür*. The QG languages also reflect **gadaa* ‘outside’, as found in Khalkha, rather than the Buriat, Khamnigan and Kalmuck innovation **gajaa* ‘id’.

The verbs **eči-*, **oči-* and **od-* ‘to go’ have a peculiar distribution. All of Shirongol except Nantoq Baoan uses **eči-*. Nantoq Baoan and Eastern Yugur use **od-* which is also found in Middle Mongol.

1.5.3. Features separating Eastern Yugur and Shirongol

1.5.3.1. Phonetic features separating Eastern Yugur and Shirongol

There are few systematic developments that separate Eastern Yugur from Shirongol. Many differences between Eastern Yugur and individual Shirongol languages are not valid classification features, because they arose after the split-up of Proto Shirongol. For instance, although Mangghuer and Dongxiang lost contrastive vowel length, it must have persisted into Proto Monguoric and Proto Baoanic, as it is still found today in Mongghul and parts of Baoan. Thus the local absence of vowel length does is not relevant for classification.

In Eastern Yugur, **ö* and **ü* did not merge with their harmonic counterparts **o* and **u*. Unlike Shirongol, Eastern Yugur does not retain traces of uncontracted diphthongs **au* and **eü*.²⁰

In Eastern Yugur the palatal affricates did not split into retroflex and alveolopalatal series.

¹⁸ The Buriat and Kalmuck forms are reminiscent of the Written Mongol spellings *eregüü*, *niruyu(n)*, *quruyu(n)*.

¹⁹ Cf. also the development of **ügtee-* ‘to pull out’.

²⁰ In Zhào (2006), and in the wordlist in Zhōng (2007) we do find diphthongs like *ɔu*. In view of other descriptions, these are best viewed as recent innovations.

The treatment of final **-n* and **ki-/ki-* is different between the two Shirongol branches, and will be discussed below. In both these matters Eastern Yugur agrees with Baoanic and differs from Monguoric.

Eastern Yugur lost the preconsonantal *-l-* in **čaalsun* ‘paper’, **mölsün* ‘ice’, and **sölsün* ‘gall bladder’ (as in Khalkha and Kalmuck), whereas Shirongol preserved it (as in Buriat and Khamnigan).²¹

Eastern Yugur tends to preserve the middle vowel of trisyllabic noun stems, whereas these were often elided (‘Mittelsilbenschwund’) in Shirongol. The words affected by this development typically have *l* or *r* preceding the elided vowel and *g* or *k* following it, e.g. **halagan* ‘palm of the hand’, **heligen* ‘liver’, **korakai* ‘insect’, **kulagai* ‘thief’. This type of elision is uncommon in Eastern Yugur, e.g. *jürŷen* < **jürüken* ‘heart’.²²

Unsystematic phonetic differences between Eastern Yugur and Shirongol include the following:

EYu	Shirongol	CM	
<i>jagčaqai</i>	<i>*čarčag</i>	<i>*čarčaakai</i>	grasshopper ²³
<i>da:lə</i>	<i>?*dalui</i>	<i>*dalu</i>	shoulderblade
<i>küre:</i>	<i>*kireü</i>	<i>*kiröe</i>	saw
<i>gertə</i>	<i>*ke(r)tü</i>	<i>*ger-tü</i>	home
<i>-kə</i>	<i>*-kü</i>	<i>*-ki</i>	‘converter’ suffix
<i>tal-</i>	<i>*tabi-</i>	<i>*talbi-</i>	to put
<i>labčəg</i>	<i>*labčün</i>	<i>*nabčün</i>	leaf
<i>hənəs-</i>	<i>*nis-, *mis-</i>	<i>*nīs-</i>	to fly
<i>aməra-</i>	<i>*ham(b)ura-</i>	<i>*hamura-</i>	to rest
<i>dörβen</i>	<i>*derben</i>	<i>*dörben</i>	four ²⁴
<i>möŋgə ~ meŋgə</i>	<i>*meŋgü</i>	<i>*möŋgün</i>	silver ²⁵

The unrounding of **ö* does not generally follow subgroup divisions like in these last examples. It often has an irregular distribution, as in BaoŃ *melsəŋ* ‘ice’ as opposed to Kgj *mərsün* < **mölsün* ‘ice’.

As to the development of **h-*, there are several other cases of disagreement between Eastern Yugur and Shirongol, like **hamura-* ‘to rest’. However, there are

²¹ These preconsonantal consonants are easily lost in modern Shirongol with its decreasing tolerance for syllable-final consonants, cf. MgrH *su:rʒə* ~ *su:ʒə* < **sölsün* ‘gall bladder’, Dgx *mansuŋ*, Dgx (Longquan dialect) *məsüŋ*, Kgj *mərsün* ~ *məsün* < **mölsün* ‘ice’.

²² The sequence *rVl* in the middle of words commonly leads to simplifications in all QG languages, as in EYu *alğa-* < **arılga-* ‘to clean’, BaoŃ *alʒa-* < **aralji-* ‘to exchange’, Kgj *kandu-* < **kereldü-* ‘to quarrel’, and even elsewhere, including Dagur, e.g. *alj-* < **aralji-* ‘to exchange’, *dull^w-* < **düreldü-* ‘to trade’.

²³ For Eastern Yugur cf. western forms like Dörbed *čagčä:xä:* (Vanduy 1965:167a).

²⁴ Only Gomar Baoan *dəraŋ* CN86:146 seems to disturb this picture, but in fact this dialect is insufficiently known to be certain how the form should be evaluated.

²⁵ The Shirongol form resembles the one in the Muqaddimat al-Adab.

also many disagreements within Shirongol. Most deviations do not neatly follow the subgroup divisions (see the words with **h-* in the comparative supplement).

1.5.3.2. Lexical and morphological features separating Eastern Yugur and Shirongol

The fact that a Mongolic word preserved in Eastern Yugur was lost in Shirongol is in itself not surprising, as the loss of native lexicon seems to be more dramatic in the latter. Nevertheless, more than a few CM words found in Shirongol were lost in Eastern Yugur as well.

Eastern Yugur lexical items missing from Shirongol include old Mongolic words, e.g. *ᠵᠣᠶᠴᠠᠢ*- ‘to sit’ < **čokai-* ‘to squat’, as well as words with a ‘Western’, i.e. Oirat, flavour, e.g. *duḡura-* ‘to roll’, *ḡaisən* ‘ghost’, *hučɯ* ‘fur jacket’, or western variants of more widely occurring words, e.g. *βəḷə* < **ubuḷi* ‘feeding horn’.

Also missing in Shirongol are ‘Yugurisms’, e.g. words shared with Western Yugur but without an obvious etymology in Mongolic or Turkic, e.g. *hanə-* ‘to go’, *lar* ‘speech, language’, *kutḡə* ‘throat’, *sure:čə* ‘crane (the bird)’, and other, possibly non-Mongolic, words apparently restricted to Eastern Yugur, such as *xalda-* ‘to look’ and *usqa-* ‘to scold’ mentioned below.

Lexical differences between Eastern Yugur and Shirongol are often the result of a different choice out of existing old vocabulary. In other cases one or both subgroups use a non-Mongolic word. Foreign words found in several Mongolic subgroups and already documented in Middle Mongol will be considered to have been present in CM.

There are numerous Mongolic lexemes that are preserved by Shirongol, but absent from Eastern Yugur. In the following cases Eastern Yugur and Shirongol use different Mongolic words. Occasionally the Yugur counterpart in this list is also attested in parts of Shirongol, but not the other way around. Note that the distribution of these same words is usually irrelevant outside the QG area, e.g. Ordos has both **elgü-* and **deiḷile-* ‘to hang’, which fact is of no classificatory value within central Mongolic.

Eastern Yugur	Shirongol	
<i>*asara-</i> , <i>*mal.la-</i>	<i>*teḷie-</i>	to feed; to raise
<i>*čiig.tü-</i>	<i>*nor-</i>	to get wet ²⁶
<i>*čokai-</i>	<i>*duṇdei-</i> (? <i>*döṇdei-</i>)	to squat, sit ²⁷
<i>*daga.ul-</i>	<i>*uduri-</i>	to lead ²⁸

²⁶ Several Shirongol idioms also use various verbs derived from the adj. **noitan* ‘wet’ with the suffixes *-rA*, *-dA*, *-tU*: BaoD *nəitar-* B85b, BaoÑ *ni:tara-* CN29, Kgj *neitɔda-* S281b, MgrM *niutiantu-* DS260a, Dgx *naićietu-* B25.

²⁷ Most of the vocabularies see the Shirongol word as derived from Ch *dūn* ‘to squat’, but as all languages involved have other common verbalizers for incorporating Chinese verbs, this may in fact be CM **döṇdei-* ‘to rise slightly’, semantically influenced by Chinese.

²⁸ **dagaul-* is the caus. of **daga-* ‘to follow’, while **uduri-* is related to the Turkic caus. of **ud-* ‘to follow’.

(continued)

Eastern Yugur	Shirongol	
* <i>dugura-</i>	* <i>höŋkeri-</i>	to roll ²⁹
* <i>elgü-</i>	* <i>deüjile-</i>	to hang
* <i>ges-</i>	* <i>simta.ra-</i>	to melt
?* <i>hamu-</i>	* <i>söne-</i>	to go out (fire)
* <i>hüsür-</i>	* <i>deül-</i>	to jump
* <i>hutaa</i> (< * <i>hutu.a</i>)	* <i>hunin</i>	smoke
* <i>kamki-</i>	* <i>hani-</i>	to close the eyes
* <i>kaurai</i>	* <i>koa.sun</i>	dry
* <i>mal</i>	(* <i>a:sun</i> <) * <i>aduu.sun</i>	livestock
* <i>möndür</i>	* <i>jarma</i>	hail
* <i>od-</i> (and <i>hanə-</i>)	* <i>eči-</i>	to go ³⁰
* <i>sura-</i>	(* <i>arsag-</i> <) * <i>harsag-</i> etc	to ask
* <i>šatu</i>	* <i>geški.ür</i>	ladder
* <i>tokanag</i> (and <i>čikeneg</i>)	* <i>tokai</i>	elbow ³¹
* <i>teü-</i>	* <i>temgü-</i>	to pick up ³²
* <i>tur-</i>	* <i>öles-</i>	to be hungry
* <i>üdesi</i>	* <i>(üde)si.leŋ</i>	(in the) evening

Eastern Yugur words of unknown origin opposed to CM words in Shirongol:

Eastern Yugur	Shirongol	
<i>xalda-</i>	* <i>kara-</i> , * <i>no-</i> , * <i>siğai-</i>	to look
<i>lar</i>	* <i>kelen</i>	language ³³
<i>ma: kelen</i> (= <i>hdei kelen</i>)	* <i>koš kelen</i>	uvula ³⁴
<i>usqa-</i>	* <i>sögee-</i> , * <i>karä-</i>	to scold, swear

The origin of Shirongol words **hayag* ‘handful’, **hurbī* ‘sling’, **ütergen* ‘threshing floor’ is unclear, but the first and third may be connected to Turkic *(*h*)*aya* ‘palm of the hand’ and **örtgün* ‘threshing floor’.

Shirongol word shapes include **aasun* < CM **aduusun* ‘livestock’.³⁵ In Shirongol the words **üdür* (?**ödür*) ‘day’ and **üde* ‘noon’ have been confused, and

²⁹ **dugura-* is not attested early, but it is also found in Kalmuck.

³⁰ EYu *hanə-* is of unknown origin, but cf. WYu *hani-*. Nantoq Baoan stands out among the Shirongol languages in that it uses **od-* rather than **eči-*.

³¹ Interestingly, **tokanag* is also found in Buriat. The similarly-structured *čikeneg* stems from Turkic.

³² Mangghuer also uses **teü-*.

³³ **kelen* does exist in Eastern Yugur in the meaning ‘tongue’.

³⁴ The meaning and origin of *ma:* is unknown. EYu *hdei kelen*, lit. ‘little tongue’ could, but need not, be a calque from Chinese; the first element is from Western Yugur. Shirongol **koš kelen* is not attested for Kangjia, which uses *šgə kelio* S84, lit. ‘little tongue’.

they generally merged into a single phonetic shape, so that a single word shape (either with or without the *-r* depending on the language) carries both meanings.

Shirongol also uses CM **kabar* ‘nose’ in the meaning ‘nasal mucus, snot’, whereas in Eastern Yugur it developed the additional meaning ‘before’ under Western Yugur influence.

A morphological innovation in Shirongol (only absent from Kangjia) involves the replacement of the habitual participle (nomen usus) *-dAG* by *-čIn*, by extending the function of the deverbal noun suffix **-g-čIn* and/or **-A-čIn*. In Shirongol the suffix *-čIn* is now used in inflection as well as word formation. Examples: MgrH *šuro:žin* ‘who blesses’, MgrM *dağačIn* ‘following; follower’, BaoŃ *medčan* ‘knowing; who knows’, Dgx *kieliečIn* ‘speaking; speaker; who speaks’. Eastern Yugur continues to use *-dAG*, and has another participle *-mA*, which it shares with Western Yugur.³⁶

Another innovation is the generalisation of the causative *-GA*. This CM suffix became invariable in Shirongol, and only the variant with uvular consonant and back vowel *-ğa* or *-ğa* remains as a productive suffix. It has replaced the causative suffixes *-Ul* and *-A*, and increasingly also *-IGA*, especially in Mangghuer, Baoan and Dongxiang. Existing formations with the old suffixes have been replaced.³⁷ Examples include MgrM *diğa-*, BaoŃ *ndeğa-*, Dgx *ižieğa-* ‘to cause to eat, feed’, MgrM *roğa-*, Dgx *oroğa-* ‘to cause to enter, insert’. MgrM *wužiğa-*, BaoŃ *užiğa-*, Kgj *užiğa-*, Dgx *užəğa-* ‘to cause to see, show’, MgrM *xuliğa-*, Dgx *fəilieğa-* ‘to cause to remain, leave over’, MgrM *bərduga-*, Dgx *barunduğa-* ‘to cause to catch fire, to light’. The Eastern Yugur equivalents are the conservative forms *edü:l-*, *ɔru:l-*, *ojü:l-*, *holü:l-*, *baldu:l-*, from the stems **ide-* ‘to eat’, **ora-* ‘to enter’, **üje-* ‘to see’, **hüle-* ‘to remain’, **bari.ldu-* ‘to catch fire’.

The loss of productive vowel harmony caused numerous other suffixes to generalise a single variant. The diminutive *-kAn* only preserved its back unrounded variant *-qan* in QG (MgrH *-xa:n*, BaoŃ *-χaŋ*, Dgx *-ğaŋ*, etc.); the nomen futuri *-kU* only preserved a form *-ku* with velar consonant.

1.5.4. Features separating Monguoric and Baoanic

1.5.4.1. Phonetic features separating Monguoric and Baoanic

There are only few systematic phonetic differences between Monguoric and Baoanic. The development of the sequence **kī-/ki-* and the treatment of unstable

³⁵ As this etymon does not survive in Eastern Yugur at all, it cannot be determined which phonetic shape it had before it was lost. It may have been **a:sun* in Eastern Yugur as well.

³⁶ EYu uses **-g-čI* in word formation, but the participial usage described in Bolčuluu & Jalsan (1992: 266) is rare.

³⁷ Several *-Ul* formations, often with specialised meanings, do survive in Shirongol, including MgrH *xarəla-*, Dgx *qarulu-* ‘to answer < cause to return’ < **kari.ul-*; BaoŃ *əčəl-*, Dgx *oçulu-* ‘to turn (tr.)’ < **horčī.ul-*; BaoŃ *ačal-* < ‘to water (animals) < cause to drink’ < **oačī.ul-*.

*-n are the most marked of these. In both these matters Baoanic is the conservative one, thus agreeing with Eastern Yugur, Moghol, and Middle Mongol.

Baoanic preserves a distinction between velar and uvular variants of */k/ in this context. In Monguoric the sequence *kī- apparently merged with its velar counterpart *ki-, which was then palatalised (> modern *ci*), before the vowel *i* lost its palatality.³⁸ Examples:

Monguoric	Baoanic	CM	
*čirğa-	*qirğa-	*kirma-	to shave
*čimusu	*qimusuŋ	*kimusun	nail
*čidoŋo	*qidoŋo	*kituga	knife

The treatment of unstable *-n is in fact a morphological matter with a phonetic result. Baoanic has incorporated unstable *-n as a permanent part of the stem. In Monguoric, unstable -n was lost in the nominative in all common nouns. Some relics of it, as well as hypercorrect *n*'s, may appear in inflected forms. That dropping the -n was morphologically motivated, is also clear from the fact that the numerals and adjectival nouns were exempted from the development. Examples:

Monguoric	Baoanic	CM	
*ima:	*imaŋ	*imaan	goat
*nudu	*nuduŋ	*nidün	eye
*toŋo:	*toŋoŋ	*togaan	pot

Most phonetic differences between Monguoric and Baoanic are irregular. In the following three forms the Baoanic languages are the innovators. Baoanic might owe the strong initial to inflected forms with suffixes starting with strong consonants, e.g. for the verb we would have Proto Shirongol future participle *gar-ku, perfect participle *gar-san, conditional converb *gar-sa.³⁹ However, such transfers of consonant strength are typically more common in Monguoric than in Baoanic. Due to its ad hoc nature this explanation is unsatisfactory. Examples:

Monguoric	Baoanic	CM	
*gar	*qar	*gar	hand
*gar-	*qar-	*gar-	to exit
*gal	*qal	*gal	fire

³⁸ The sequence *kī/*ki in medial position, and its weak counterpart *gi/*gi shows similar developments.

³⁹ For *gar and *gar- there would be the additional factor that some variable suffixes (e.g. imperfect converb *-ŋU, dat. *-DU) used to take the strong consonant variant after stems in -r). Stems in -l took the weak consonant variant, so why *gal 'fire' developed a strong initial would remain unexplained.

Possibly cases like these should simply be counted among the numerous irregular correspondences regarding consonant strength, especially in words with initial **k-* and **g-*. There are two known environments than can affect the consonant strength of these consonants in all of QG. Strengthening of **g-* may occur when the second syllable starts with a strong consonant, while weakening of **k-* commonly occurs when the second syllable starts with **d* or **j̣*. Strengthening or weakening in the absence of these known triggers, or unexpected preservation of strong or weak consonants, may also create differences between Monguoric and Baoanic. E.g. CM **kura* ‘rain’ developed into expected **qura* in Monguoric, but underwent weakening in Baoanic **gura*. Likewise CM **mikan* ‘meat’ resulted in Monguoric **maqa*, but Baoanic **mağa*.

Unsystematic phonetic differences between Monguoric and Baoanic include the following (either side can be the innovative or irregular one):

Monguoric	Baoanic	CM	
<i>*dabsağ</i>	<i>*dabalağ</i>	<i>*dabusag</i> , etc.	bladder ⁴⁰
<i>*gedesun</i>	<i>*gejesun</i>	<i>*gedesün</i>	intestine
<i>*uŋgasun</i>	<i>*noğosun</i> < <i>*nogasun</i>	? <i>*nuŋgasun</i>	wool
<i>*su:</i>	<i>*suğu</i>	<i>*suu</i> (~ <i>*sugu</i>)	armpit
<i>*qarmu-</i>	<i>*qarbu-</i>	<i>*karbu-</i>	to shoot
<i>*ta:da</i>	<i>*ča:da</i>	<i>*čaada</i>	near ⁴¹
<i>*qaučin</i>	<i>*qoičin</i>	<i>*kaučin</i>	old
<i>*aur</i>	<i>*hor</i>	<i>*haur</i>	air; anger
<i>*janči-</i>	<i>*yanči-</i>	<i>*janči-</i>	to beat, to thresh ⁴²

1.5.4.2. Lexical and morphological features separating Monguoric and Baoanic

Some differences between Monguoric and Baoanic are due to morphological innovation. **yamar* ‘what kind?’ was retained in Monguoric, but replaced in Baoanic by **yaan metü* or **yama metü*, lit. ‘what-like?’. In Baoanic the CM words **eimü* ‘this kind of’, **teimü* ‘that kind of’ were largely replaced by similar formations **ene metü* ‘this-like’, **te(re) metü* ‘that-like’, or possibly **ein metü* and **tein metü* ‘thus-like’.⁴³ The degree of phonetic reduction in the modern languages makes it hard to distinguish the exact components, cf. Baoñ *təmtəg* (the final *-g* appears to be secondary), Kgj *tentü*, Dgx *ćimutu* ‘that kind of’. For further forms see the comparative supplement s.v. **metü*. In Monguoric, **eimü* and **teimü* were replaced

⁴⁰ The Monguoric and Baoanic forms may represent different derivatives of the same root, but the endings are not clear.

⁴¹ Cf. Mongghul *tagṣa* ~ *čagṣa* < **čaa-gsī* ‘to(wards) that side’.

⁴² The Baoanic form with *y-* is reminiscent of the Turkic cognate **yanč-*. See **jarim* ‘half’ for a similar case.

⁴³ The latter etymology is favoured by Dongxiang *ćimutu* ‘that kind’ which is easier to explain with *ćin* < **tein* as a first element than with **tere* which (irregularly) retained its *t-* in Dongxiang.

by other formations of unclear structure: Mongghul *nəgi*: ‘this kind of’, *təgi*: ‘that kind of’, Mangghuer *niŋtai* ‘this kind of’, *tiŋtai* ‘that kind of’.

The decads higher than ‘twenty’, such as **gučīn* ‘thirty’, **döčīn* ‘forty’, were replaced in Baoanic by analytical forms **gurban harban* ‘three tens’, **dörben harban* ‘four tens’.⁴⁴ In Kangjia these were retained as analytical forms *guro haro*, *dero haro*, but in all Baoan dialects they have started to amalgamate, e.g. Nantoq Baoan *gu‘raraŋ*, *de‘raraŋ*. Kangjia even replaced **korīn* ‘twenty’ by an analytical form ‘two tens’. Dongxiang preserves *qoruŋ* < **korīn* ‘twenty’, but has now replaced all higher decads by Chinese forms. It is impossible to say if the analytical Mongolic forms were ever used in Dongxiang.

In Baoan and Dongxiang the numerals 1-10 tend to join with certain nouns that follow them.⁴⁵ This phenomenon occurs with the native words **üdü* ‘day’ and **ayaga* ‘bowl’, and a number of quantifiers and classifiers.⁴⁶ Some of the resulting forms: Dgx *žierudu* (for *žieroŋ udu*) < **dörben üdü* ‘four days’, *hariğa* (for *haroŋ iğa*) < **harban ayaga* ‘ten cups’, BaoD *de‘rudə* (for *derəŋ udər*) < **dörben üdü* ‘four days’, *har‘wiğə* (for *harwaŋ ayiğə*) < **harban ayaga* ‘ten cups’.⁴⁷ As the abovementioned nouns are often used in combination with numerals, it is understandable that they are worn down in this manner. However, it is unclear why other commonly counted nouns like **söni* ‘night’, **nasun* ‘year (of age)’ and **küün* ‘person’, should escape this treatment. It may be due to the fact that these words, unlike **üdü* and **ayaga*, have initial consonants, but perhaps more importantly, **üdü* ‘day’ and **ayaga* ‘bowl’ are also used locally in the meaning ‘thousand’ and ‘ten thousand’, respectively.⁴⁸ It cannot be established whether such joined forms ever occurred in Kangjia, as it systematically lost the final *-n* of these numerals.⁴⁹

In other cases different derivations of the same stem prevailed in the two Shirongol groups. In Baoanic the spatial noun **emüne* ‘front’ is retained in the form **emüle* (with dissimilation of the nasals), whereas in Monguoric a shortened form of the directive **emüne-gsi* ‘to the front’ has replaced its base.⁵⁰

⁴⁴ The forms replacing the old decads are transparent and obvious, and need not have been inspired by Chinese, Tibetan or Turkic, where similar forms exist.

⁴⁵ In Baoan, as in Tibetan and other languages influenced by it, numerals may also follow the noun, in which case this ‘joining’ does not take place.

⁴⁶ These may be of Chinese, Tibetan, or unknown etymology, e.g. Dgx *fa* ‘times’, *suay* ‘pair’, BaoÑ *yaŋ*, *taŋ*, *gə* ‘times’. See Bökh & Čoyijungjab (1985:114) and Chén Nǎixióng et al. (1986:163) for Dongxiang and Baoan examples.

⁴⁷ The numerals lose their *-n* in the process. In Dongxiang, *gua* < **koar* ‘two’ regains the *-r* it normally lacks, as in *guariğa* ‘two cups’, *guarudu* ‘two days’.

⁴⁸ The additional meanings were calqued via Chinese. See Nugteren & Roos (2010).

⁴⁹ According to Li Keyu **üdü* also occurs in compounds with some numerals in Mongghul: *go:du* ‘two days’, *guru:du* ‘three days’, *xaru:du* ‘ten days’, *xuru:du* ‘twenty days’ (from **koar*, **gurban*, **harban*, **korīn*). Li also mentions forms with **söni* ‘night’: *xuru:sonə* ‘twenty nights’, *xužu:sonə* ‘thirty nights’ (**korīn*, **gučīn*). These forms appear to have developed independently from Baoanic, as **ayaga* and **üdü* did not develop numeral functions in Mongghul.

⁵⁰ MgrH *mušī*, MgrM *mešī*. As this form lost its specific function, a new directive form was created: MgrH *mušīžə*, MgrM *mešīžī*.

Due to the loss of vowel length in most of Baoanic and in Mangghuer, and accent-related vowel lengthening in Mongghul several verbs were at risk of merging with their causatives in -A, resulting in a need to make the distinction clear again. The chosen solution may vary from verb to verb. In case of **sita-* ‘to catch fire’ and its causative **sītaa-* ‘to make fire’, Baoanic created a new intransitive verb **sītara-* with a new causative **sītaraga-*, whereas in Monguoric, as in Eastern Yugur, only the transitive **sītaa-* ‘to light’ survives.

The ‘converter’ suffix **-kl*, which turns an inflected noun into a new nominative base that can take case endings, split into two forms in Baoan and Dongxiang, whereas in Monguoric a single form is used.

Attached to stems with the dative, as well as to adverbs of time without explicit case marking, we find the Shirongol shape *-ku* with unexpected rounding, as in Dgx *araŋdaku* ‘the one in the river’, *məliešidaku* ‘the one in front’, *fuzuğuduku* ‘yesterday’s’, BaoŃ *kətagu* ‘the one in the house’, *ğadəgu* ‘the one on the outside’, *çi:nagu* ‘the one in back’, *maχsīgu* ‘tomorrow’s’.

To stems with the genitive a form with *-ŋ* is attached, perhaps from an old oblique form **-kIn-*, as in Dgx *minuğun*, BaoŃ *məŋgaŋ* ‘mine’. This can be found with the old personal pronouns, but also with the non-CM pronouns such as Dgx *hənuğun*, BaoŃ *aʒaŋgaŋ* ‘his/hers’, and some other pronouns and nouns, e.g. Dgx *kienuğun* ‘whose’, *gienuğun* ‘the house’s’, BaoŃ *kəŋgaŋ* ‘whose’, *ndewağgaŋ* ‘that of the village’.⁵¹

In Kangjia the suffix shape *-gu* is used on genitives as well as datives, e.g. *məniɡu* ‘mine’, *kəɡu* ~ *kəniɡu* ‘whose’, *ɡanaɡu* ‘the one on the outside’, *kəŋɡu* ‘the one of the person’, *ağagu* ‘the one of the village’, *ağadugu* ‘the one in the village’.⁵² This situation is reminiscent of the central Mongolic situation (e.g. Ordos *geri:ki* ‘the one of the house’, *gerteki* ‘the one in the house’). In Monguoric and Eastern Yugur the genitives of personal pronouns seem to be unable to take the converter suffix, that is, the attributive shape of the genitive is also used predicatively.

Lexical differences come in various categories. In the first set below a different CM word has prevailed in Monguoric and Baoanic. In the second group, Monguoric replaced a well-known Mongolic word by one of unknown origin. In the third group it is Baoanic that has the replacement, of Turkic or unknown origin.

Monguoric	Baoanic	CM	
<i>*idee</i>	<i>*höesün</i>		pus
<i>*muruï</i>	<i>*maïrug</i>		crooked
<i>*üğtee-</i>	<i>*sei-</i>		to weed
<i>*anji:</i>	<i>*qala</i>	<i>*kaana</i>	where ⁵³
<i>*baga-</i>	<i>*eki-</i>	(various) <i>*högi-</i>	to hit ⁵⁴

⁵¹ *aʒaŋgaŋ* (for **aʒaŋnəgaŋ*) lacks the genitive suffix. Other commonly used nouns such as ‘house’ and ‘village’ also add *-ku* directly to the stem without the need for a genitive.

⁵² This may also be the case in some Baoan dialects, cf. BaoD *ćin(ə)gə*, BaoX *ćinəgo* ‘yours’ as opposed to BaoŃ *ćinɡaŋ*.

⁵³ Kgj *χana* deviates from its relatives here.

Monguoric	Baoanic	from	
* <i>silgüd-</i>	* <i>taqara-</i>		to shake
* <i>čaalsun</i>	<i>kegde</i>	< T < Ir	paper
* <i>šu:r</i> ⁵⁵	<i>čöke</i>	< T	chopsticks
* <i>balgasun</i>	<i>tam</i>	< T	wall

Some lexical differences are due to semantic shifts. CM **ejen* ‘master’, as in Eastern Yugur, has become a reflexive pronoun ‘self’ in Monguoric, whereas in Baoan (not in Kangjia and Dongxiang) it became a personal pronoun for the third person. CM **öer-* ‘self’ was retained in Baoanic as a personal pronoun (denoting the first person in Baoan and Dongxiang, and the third person in Kangjia), but was lost in Monguoric and Eastern Yugur.

CM **jaka* ‘collar’ developed the additional meaning ‘upside, top’ in Monguoric (also as in Eastern Yugur).⁵⁶ This probably coincided with the loss of **öede* ‘up, above’, which does survive in Baoanic. **no-* (?**nau-*) ‘to aim’ came to mean ‘to look’ in Monguoric.

As seen above, certain loanwords, especially those from Turkic, occur in informative patterns that support the subdivision of Shirongol. The Monguoric languages share **arčag* ← Turkic **agırçak* ‘spindle whorl’ and **tagau* ← Turkic **takagu* ‘chicken’. Baoanic shares **örüg* ← Turkic **ärük* ‘apricot’, **kegde* ← Turkic (← Iranian) **kägdä* ‘paper’, and **sičagan* ← Turkic **sičgan* ‘mouse’. In some instances Shirongol languages replaced a lexeme by its Turkic cognate. Baoanic borrowed the Turkic word and lost the native cognate in the following cases: Turkic **agil* ‘village’, **arpa* ‘barley’ and **bugday* ‘wheat’ are used instead of CM **aïl*, **arbai* and **buudaï*. In Monguoric, Turkic **takagu* ‘chicken’ has replaced native **takia*.

1.5.5. Deviating distribution of word shapes

1.5.5.1. Introduction

On several occasions, the data from the QG languages do not align with the patterns and subdivisions sketched above.

First of all, some data seem too erratic to be organised in a way that supports (or contradicts) the classification above, such as the verb ‘to read’: EYu *oŋšə-*, MgrH *mošə-*, MgrM *moši-*, BaoD *muši-*, BaoGt *məš-*, BaoÑ *əmci-*, Kgj *unči-*, Dgx *oŋši-*. Another lexeme whose modern forms seem to defy classification (or reconstruction) is ‘girl’: EYu *hkon*, MgrH *fužün*, MgrM *fužin*, BaoD *okən*, BaoÑ *okuŋ*, BaoX *xjoŋ*, Kgj *iχgə*, Dgx *očin*. Cases like this may be explained by inter-

⁵⁴ Kangjia has both *iŋ-* related to the Baoanic verb and a form *bīŋi-* apparently related to Monguoric.

⁵⁵ According to Sun, EYu *šoro* < **širo* ‘skewer’ (unrelated to the Monguoric word) is also used for ‘chopsticks’.

⁵⁶ Kangjia *jīgali* ‘top’ may also be related, but is morphologically unclear. In Western Yugur, the Turkic equivalent **yaka* ‘collar’ is now also used as a spatial noun.

dialectal borrowing, as is suggested by the existence of doublets in individual languages, e.g. Mongghul *ḡusḡo-* and *ṣdoḡo-* ‘to stir’ seem to be divergent developments of **kudku-*.⁵⁷ Mongghul *ḡazar* and *ḡadar* ‘bridle’ seem to represent the old variants **kaḡaar* and **kadaar*. Mongghul *go:n*, *gom*, and *fugon* are divergent forms of **gün* ‘deep’. The verb **bagta-* ‘to sink’ is present in Kangjia twice, in the forms *paḡda-* and *puda-* (with semantic differentiation). It is difficult to distinguish such doublets from the ‘normal’ spectrum of variants encountered in these non-standardised languages, and to determine which of the forms was borrowed from another language.⁵⁸

Semantically similar lexemes may have an erratic, and thus uninformative distribution, such as the words for ‘head’ and ‘brain’: **hekin*, **taraki*, **teriün*, **tolagai*.

Another factor that disrupts the classification as discussed above, is retention of shared inherited features, e.g. Eastern Yugur agrees with Baoanic regarding the development of **kī-/ki-* and **-n*, which is irrelevant for classification, whereas the innovations found in Monguoric can be used to define that branch.

Finally there are many similarities that can be explained as *Sprachbund* phenomena as well as shared inherited features.

There are three recurrent deviations from the subdivisions described above:

- Eastern Yugur agreeing with Monguoric (or at least Mongghul)
- Mangghuer agreeing with Baoanic rather than with its sister language Mongghul
- Disagreements amongst the Baoanic languages

1.5.5.2. Similarities between Eastern Yugur and Monguoric

In the following cases Eastern Yugur and Monguoric share a phonetic variant:

E. Yugur	Mongghul	Mangghuer	CM (or local ancestral form)	
<i>ḡḡwa:sən</i>	<i>ḡḡua:sə</i>	<i>ḡḡuasī</i>	<i>*uḡgasun</i>	wool
<i>ḡö:rö ~ ḡüre</i>	<i>go:ro</i>	<i>gori</i>	<i>*öere</i>	other

In the first case Baoanic uses the related variant **nogasun*. **öere* ‘other’ was lost in Baoanic, but cf. **öermiče*. Eastern Yugur and Monguoric preserve **bida* ‘we’, while Baoanic seems to feature the innovative front-vocalic form **bide*. This certainly applies to Dgx *biḡien* and Kgj *bəde*. Most of the Baoan forms are ambiguous because of vowel reductions, but BaoGt *buda* stands out by agreeing with Eastern Yugur and Monguoric.

Eastern Yugur and Monguoric, or at least Mongghul⁵⁹, share the following lexemes that are absent from Baoanic:

⁵⁷ Cf. **kadku-* ‘to stab’ for similar Mongghul doublets.

⁵⁸ For a Dagur doublet see *čulč* and *sus* under **sölsün* ‘gall bladder’.

⁵⁹ The cases not confirmed by Mangghuer may be due to the larger number of loanwords in that language, to the smaller size of the wordlists, or to the fact that Mangghuer sometimes agrees with Baoanic rather than with its sister language.

E. Yugur	Mongghul	Mangghuer	local ancestral form	
<i>mənən</i>	<i>muno:n</i>	---	<i>*mona hon</i>	year after next
<i>gərβai</i>	<i>ćire:</i>	<i>çaibai</i>	<i>*kirbei</i>	edge
<i>də:mbər</i>	<i>dumbur</i>	---	?	hill ⁶⁰
<i>ǰura-</i>	<i>ǰira:-</i>	<i>zua-</i>	<i>?*ǰura-</i>	to chase
<i>χərǰər</i>	<i>śorǰol</i>	---	<i>?*hürgul</i>	omen ⁶¹
<i>səis</i>	<i>śiās</i>	---	<i>?*süis</i>	billy-goat
<i>boda</i>	<i>bodo</i>	---	<i>*boda</i>	deer ⁶²

1.5.5.3. Similarities between Mangghuer and Baoanic

In some phonetic developments Mangghuer agrees with Baoanic rather than Mongghul, i.e., Mongghul stands out within Shirongol. In most of the following cases, Mongghul is the conservative language, lacking the irregular development shared by the other languages. Thus Mongghul agrees with both Eastern Yugur and central Mongolic in those cases.

Mongghul	Mangghuer	Baoanic	CM	
<i>fod</i> < <i>*hod</i>	<i>xotu</i>	<i>*hötüg</i>	<i>?*hötü</i>	maggot
<i>nes-</i>	<i>musi-</i>	<i>*mis-</i>	<i>*nis-</i>	to fly
<i>tailə-</i>	<i>tai-</i>	<i>*tai-</i>	<i>*tail-</i>	to untie
<i>budən</i>	<i>beduŋ</i>	<i>*bedüŋ</i>	<i>*bidüün</i>	coarse
<i>ćidar</i>	<i>ćidai</i>	<i>*kitai</i>	<i>*kitad</i>	Chinese

Strictly speaking the form **kitai* in the last example is not an erratic phonetic development but a shared loanword from Turkic.

In case of **seül* ‘tail; end’, the Baoanic forms stem from an irregularly unrounded **sel*, which also seems to be the source of MgrM *śier*.⁶³

A difficult case is **yeke* ‘large’, which developed rounded vowels in the Baoanic languages, as if there was an intermediate form **hüke*, cf. BaoD *fgo* ~ *go*, BaoÑ *ǰgə*, BaoX *xgo*, Kgj *gɬ*, Dgx *fugie*. Rounding is also found in MgrM: *ǰ(u)go*, but not in MgrH: *ǰge*. However, it is unclear what triggered the rounding, and what caused the divergent modern forms.⁶⁴

⁶⁰ BaoÑ *dəmbəg* ‘id’ may be (irregularly) related. Cf. also Sanchuan (MgrM) *tomlùk* P416b ‘ancient (?grave) mound’.

⁶¹ Perhaps somehow related to **iro*, cf. the extended stem in Kalm *yorǰo.ći* R219b ‘soothsayer’. However, the absence of **h-* in **iro* is an argument against such a connection. The enigmatic Middle Mongol form (Ibn Muhanna) *hilyab* (P438a) ‘omen’ may represent a scribal error of the same etymon.

⁶² This may the same word as **boda* ‘large livestock’.

⁶³ However, Mangghuer forms with preserved labial element are also recorded.

⁶⁴ The presence of the initial fricatives (normally suggestive of initial **h-*) is also unexpected. It could be the result of devoicing of the first syllable by the following *-k-*. However, the Dagur cognate *xiy* may support the presence of **h-*.

The unexpected strengthening of **-d-* seen in MgrM *xotu* ‘larva’ mentioned above is also present in several other words, with unsystematic parallels in the Baoanic languages. It cannot be excluded that Mongghul originally shared this phenomenon, and lost it more recently due to newly developed restrictions on the distribution of consonant strength.

Mangghuer	Mongghul	Ñantoq Baoan	CM	
<i>mutu</i>	<i>mo:də</i>	<i>mətəŋ ~ mədəŋ</i>	<i>*modun</i>	wood
<i>xotu</i>	<i>fo:də</i>	<i>hətəŋ</i>	<i>*hodun</i>	star

Strengthening of the final **-d* of verbs can also be observed.⁶⁵ This involves the metanalysis of the connective vowel *U* as a part of the stem. The many verb suffixes with strong consonants may have triggered the strengthening (**-d-* that was already intervocalic in CM seems to escape this development).

Mangghuer	Mongghul	Dahejia Baoan	CM	
<i>betu-</i>	<i>idə-/udə-</i>	<i>etə-</i>	<i>*ebed-</i>	to hurt
<i>qaputu-</i>	<i>xai:də-</i>	<i>χitə-</i>	<i>*kabid-</i>	to swell

Lexical differences between the two Monguor languages are usually not informative. Many lexemes that are found in Mangghuer but not in Mongghul can either be well-known Mongolic words such as *bečín* < **ebečín* ‘disease’, *qudaŋ* < **kotan* ‘enclosure’, *kuardi-* < **kölde-* ‘to freeze’, *nažər* < **nažir* ‘summer’, or unique words of unknown origin such as *dama* ‘face’ and *duguli* ‘demon’.

**yama* ‘thing’ has developed the meaning ‘food’ in Mangghuer, as it has in Baoanic. This is paralleled by Turkic **neme* ‘thing’ in neighbouring Salar.

A functional development shared by Mangghuer and Baoanic is the use of the collective numeral **koarla* (< **koar-ula*, replacing CM **koya-ula*) ‘the two of us/you/them’ as a new comitative and instrumental suffix,⁶⁶ generally in a simplified shape: MgrM *-gula*, *-guerlo*, BaoGm *-guala*, BaoGt *-galə*, BaoÑ *-gala*, Kgj *-gala*, Dgx *-gala*.⁶⁷

1.5.5.4. Differences within Baoanic

Each of the Baoanic languages Baoan, Kangjia and Dongxiang has its own defining innovations not found in the other two subgroup members. Dongxiang developed *ʒ* and *č* from **d* and **t* preceding certain vowels (inspired by local Chinese). Baoan has the replacement of initial high vowels (followed by **d* or **j*) by *n-*. Kangjia has the development word-final *-ɔ* (via *-ō*) from **-an* and **-en*.

⁶⁵ Also in Moghol, see 4.3.3.

⁶⁶ The old comitative/instrumental suffix *-la* < **-IUA* has not disappeared from Shirongol. In Mangghuer, Kangjia, and Dongxiang it is still attested in both functions.

⁶⁷ In some dialects a less abraded form is preserved in the numeral function, as in BaoÑ *guala*, Dgx *guala* ‘the two of us/you/them’.

BaoÑ	Kgj	Dgx	CM	
<i>te:r-</i>	<i>tør-</i>	<i>ćiauru-</i>	<i>*teberi-</i>	to embrace
<i>dulə-</i>	<i>dauli-</i>	<i>źiauli-</i>	<i>*deüli-</i>	to jump
<i>fulde-</i>	<i>hundi-</i>	<i>fəñźie-</i>	<i>*hülde-</i>	to chase
<i>ndaŋ</i>	<i>idə</i>	<i>uiźien</i>	<i>*eüden</i>	door
<i>nde-</i>	<i>ide-</i>	<i>iźie-</i>	<i>*ide-</i>	to eat
<i>nda:-</i>	<i>uda-</i>	<i>uda-</i>	<i>*uda-</i>	to be late
<i>tarəŋ</i>	<i>tarə</i>	<i>tarəŋ</i>	<i>*tarian</i>	crop
<i>telaŋ</i>	<i>čilə</i>	<i>tulien</i>	<i>*tüljen</i>	firewood
<i>altaŋ</i>	<i>antə</i>	<i>antaŋ</i>	<i>*altan</i>	gold
<i>źulaŋ</i>	<i>žilə</i>	<i>zəlien</i>	<i>*jöelen</i>	soft

These typical developments only have sporadic equivalents in the related languages. For instance the Baoan ‘prenasalisation’ is found once in Kangjia in *ndasun* ~ *dasun* < **hutasun* ‘thread’. The Dongxiang type of palatalisation is also occasionally found elsewhere, e.g. BaoÑ *ćiźe-* and Kgj *ćiźe-* agree with Dgx *ćieźə-* < **teźie-* ‘to feed’. An early instance of this phenomenon (as shown by the fact that Dongxiang further developed the palatal into a retroflex) is seen in BaoD *gaźiasuŋ*, Kgj *gaźisun*, Dgx *kizəsui* < **gedesün* ‘bowels’.

In the following cases Baoan has contracted the diphthongs that were preserved by Kangjia and Dongxiang:

BaoÑ	Kgj	Dgx	CM	
<i>kitaŋ</i>	<i>kuitə</i>	<i>kuićien</i>	<i>*köiten</i>	cold
<i>itaŋ</i> (BaoD)	<i>uitə</i>	<i>uitaŋ</i>	<i>*huićan</i>	narrow
<i>çi:ćila-</i>	<i>qaićila-</i>	<i>qaićila-</i>	<i>*kaići.la-</i>	to clip

A conservative feature of Kangjia is the preservation of four rounded vowel qualities. Although several forms are diachronically surprising, the following forms illustrate some of the vowel distinctions that were lost in Baoan and Dongxiang.

BaoD	Kgj	Dgx	CM	
<i>hor</i>	<i>hər</i>	<i>xo</i>	<i>*haur</i>	steam; anger
<i>hor</i>	<i>hər</i>	<i>xo</i>	<i>*heür</i>	nest
<i>tosuŋ</i>	<i>tusun</i>	<i>tosuŋ</i>	<i>*tosun</i>	fat, oil
<i>bosuŋ</i>	<i>bəsun</i>	<i>bosuŋ</i>	<i>*böesün</i>	louse
<i>hosuŋ</i>	<i>həsun</i>	<i>hosuŋ</i>	<i>*höesün</i>	pus
<i>gurdun</i>	<i>gurdun</i>	<i>gudun</i>	<i>*kurdun</i>	fast
<i>nəduŋ</i>	<i>nudu</i>	<i>nuduŋ</i>	<i>*nidün</i>	eye

No classificatory information can be extracted when each language shows a different development, as in the case of word-final **-l*, where Baoan is the conservative one, and Kangjia and Dongxiang have different alterations neither of which can be derived from the other.

BaoÑ	Kgj	Dgx	CM	
<i>kəl</i>	<i>kuar ~ kər</i>	<i>koŋ</i>	<i>*köl</i>	foot
<i>χal</i>	<i>χar</i>	<i>qaŋ</i>	<i>*gal</i>	fire
<i>sel</i>	<i>cer</i>	<i>śien</i>	<i>*seül</i>	tail

Kgj *χar* < **gar* ‘hand’ and *χar* < **gal* ‘fire’, as opposed to Dgx *qa* < **gar* ‘hand’ and *qaŋ* < **gal* ‘fire’, also demonstrate that the Kangjia situation cannot have developed from Dongxiang (in which case the Kangjia forms could not be identical), nor vice versa (in which case Dongxiang would have the form *qa* for both words).

Another matter in which the three Baoanic languages show different reflexes is the treatment of the vowels *e* and *a* after *ö* or *o* in the preceding syllable. Kangjia has typically rounded the *e* and *a*, and then raised both vowels. Interestingly, Dongxiang often shows an unrounded second syllable in front-vocalic words, and Baoan in back-vocalic words. In view of other words in Dongxiang and Baoan that do feature the expected rounding, and equivalents in other languages, the unrounded forms are probably secondary developments from Proto Shirongol forms with labial harmony.⁶⁸

BaoÑ	Kgj	Dgx	CM	
<i>kugə</i>	<i>kuku</i>	<i>kugie</i>	<i>*köke</i>	blue
<i>nəχgər</i>	<i>nəχgu</i>	<i>nokie</i>	<i>*nöker</i>	friend
---	<i>sugu-</i>	<i>sugie-</i>	<i>*sögee-</i>	to scold
<i>dəgla-</i>	<i>duḡlu-</i>	<i>doḡolo-</i>	<i>*dogal-</i>	to limp
<i>ʂdəgə</i>	<i>duḡu</i>	<i>qudoḡo</i>	<i>*kituga</i>	knife
<i>ələŋ</i>	<i>ulu</i>	<i>oloŋ</i>	<i>*olan</i>	much
<i>χələ</i>	<i>ḡulu</i>	<i>ḡolo</i>	<i>*kola</i>	far
<i>səlḡa</i>	<i>sulḡu</i>	<i>soŋḡo</i>	<i>*saulga</i>	bucket

The abovementioned features, shared by different sets of Baoanic languages, do not bring us closer to a binary division.

Moreover, a wide variety of reflexes may be observed within a single language as well, e.g. the treatment of word-final **-g* in Kangjia: elision in *ima* < **aĩmag* ‘village’, a fricative in *aləḡ* < **alag* ‘variegated’, and a plosive in *pəḷəḡ* < **burčag* ‘bean’ (cf. 4.5.5.).

Some lexemes or variants in Kangjia differ from Baoan and Dongxiang, or even from all other QG languages. Examples include Kgj *čürsün* ~ *čürsa* < **čölsün*

⁶⁸ Some of the words in question, including **kituga* and **saulga* listed here, do not display labial harmony in central Mongolic, as they originally did not have **o* in the first syllable.

‘gall bladder’ instead of the variant **sölsün* found in the other Shirongol languages, *gudər* < **gutul* ‘boots’ instead of the variant **gutusun* prevailing in the other QG languages, and *χana* < **kaana* ‘where’ instead of the innovative form **kala* found in Baoan and Dongxiang.⁶⁹

In view of the above I will provisionally assume that Baoan, Kangjia, and Dongxiang are three equal subdivisions of Baoanic.

⁶⁹ This form has an equivalent in Salar *ġala* ‘where’; the structure of the forms with *-l-* is unclear in all languages involved.