

Peter Schrijver

13 East Caucasian perspectives on the origin of the word ‘camel’ and some notes on European substrate lexemes

1 A brief history of *camel*

The word family to which English *camel* belongs is nowadays widespread in all Germanic and Romance languages. Its immediate origin is Latin *camēlus*, which in turn was borrowed from Greek κάμηλος. The latter is attested since the early fifth c. BCE (Aeschylus, Herodotus; Chantraine 1968: 489). The Greek term is generally assumed to have been borrowed from West Semitic (Masson 1967: 66): compare e.g. Biblical Hebrew *gāmāl*, Syriac *gamlā*, Sabeian *gml*, Arabic *ṣam(a)l* (Militarev & Kogan 2005: 116–118).

2 Formal problems facing the idea that Greek was borrowed directly from Semitic

Borrowing is entirely as expected of the name of an animal that is not native to Europe nor traditionally used there as a domesticated animal. In this particular case, the Greek form lies at the base of all related forms in European languages. The hypothesis that the Greek form was itself borrowed from Semitic requires two formal steps: the *-ā-* of the second syllable in Semitic (Hebrew) *gāmāl* was adopted into Ionian Greek, where it joined inherited **-ā-* in becoming *-ē-* (Kretschmer 1892: 287). Apparently the first *-ā-* was not so treated, however, for reasons that remain unclear. Furthermore, the initial *g-* of Semitic was replaced by Greek *k-*. This presents a serious problem (Heide 2011: 363; Militarev & Kogan 2005: 118): Greek possesses in its phonemic inventory a voiced *g-* like Semitic does, so there appears to be no reason to regard the substitution of *g-* by *k-* as an adaptation to Greek phonology. The *g* does make an isolated appearance in the compound name Γαυγάμηλα = καμήλου οἶκος ‘camel house’ (Strabo 16, 1, 3), but apart from this single instance in what is clearly a non-Greek name, Greek consistently shows *k-*. Therefore, whereas there can be no doubt that Greek κάμηλος was borrowed and that it is related to the Semitic word family, the attested Semitic forms may not be the direct source from which it was borrowed. The word for camel is not attested in Phoenician (apart from the word’s use as the name of the third letter of the alphabet, underlying Greek *gamma*), which

would be a good candidate for being the immediate source for historical reasons, but if the Phoenician word had initial *k-* rather than expected *g-*, that would have been irregular and in need of an explanation itself. Benjamin Suchard and Maarten Kossmann point out to me that Phoenician is an unlikely source of the Greek form for another reason as well: since long *ā* regularly became long *ō* in Phoenician, the Greek word for ‘camel’ should have had *ō* rather than *ā* > *ē* in the second syllable (cf. the letter name *yād* > Phoenician *yōd* > Greek *iota*).

According to Yakubovich (2016: 87 with footnote 30),¹ Greek *k-* may be explained if the Semitic word was borrowed into Greek via the intermediary of an Anatolian language, where word-initial voiced plosives do not occur and are replaced by voiceless plosives. He suggests that Hieroglyphic Luwian *ka-mara/i-* may mean ‘camel’ (which is a guess, cf. Bauer 2020). A borrowing via Anatolian may indeed account for the Greek *k-*, but even if Hieroglyphic Luwian *ka-mara/i-* means ‘camel’, it is unlikely to have been the immediate source of the Greek word because it accounts for neither the Greek *-ē-* nor the Greek *-l-*.

In this contribution, an alternative will be presented, viz. that the Greek term may have been borrowed from a language that is cognate with East Caucasian (see section 5).

3 Bactrian camels, dromedaries and their domestication

The Semitic world is an obvious area to which one might look for the origin of a word for ‘camel’, but it is not the only conceivable area. The Greek term (Chantraine 1968: 489), like its Semitic counterparts (Heide 2011: 363), is a generic word for ‘camel’, which denotes both the two-humped Bactrian camel (*Camelus bactrianus*) and the single-humped dromedary (*Camelus dromedarius*). Nowadays, the Bactrian camel is a rare native of the cold deserts of Central Asia, but its original territory probably included wider stretches of the Central Asian steppe. The earliest known traces of domesticated Bactrian camels are from Turkmenistan and date from the middle of the 3rd millennium BCE. In Namazga IV levels at the urbanized site of Altyn Depe, a clay model of a Bactrian camel drawing a cart was discovered which dates from ca. 2400 BCE (see recently Kirtcho 2009; Mallory & Adams 1997: 389, also 135–136). This early Bronze Age culture, which probably separated Indo-European pastoralists living on the steppe from more southerly

¹ I am indebted to Zsolt Simon for drawing my attention to this idea and to Yakubovich’s article.

regions that correspond to present-day Iran, Pakistan and India, probably spoke a non-Indo-European language. Slightly later, between ca. 2250 and 1700 BCE, it was replaced by the Bactria Margiana Archaeological Complex (BMAC), which plays an important part in the migration of Indo-Iranian speakers from the steppe to the south (Mallory in Mallory & Adams 1997: 72–74). Camelid skeletal remains in quantity appear around 3000 BCE in southern Central Asia and in eastern Iran (Kohl 2007: 142–143). By the third millennium BCE, the domesticated camel probably opened up access from the east towards the western steppe north of the Caucasus because it made regular travel across the desert steppes east of the Caspian possible (Kohl 2007: 62, 200, 250). Camel remains have been found in western steppe Yamnaya sites (3600–2200 BCE; Mallory and Adams 1997: 651).

Indirect evidence for a much earlier date of the domestication of the Bactrian camel derives from the Xinglongwa culture of Inner Mongolia, which flourished ca. 6100–5300 BCE: its desert sites lie so far apart across inhospitable terrain that people must have used camels to travel between them (Potts 2004: 148).

Very soon after the earliest evidence for domesticated camels in Turkmenistan they make an appearance in Mesopotamia. The earliest record of the Bactrian camel is in an animal list from the mid-third millennium BCE: Sumerian **am.si.ḫar.an**, literally ‘elephant of the road/caravan’. A different Sumerogram is **am.si.kur.ra**, literally ‘elephant of the land/mountains’. Both are transparent neologisms (Heide 2011: 354–60).

Wild populations of the dromedary do not exist anymore. It is presumed to have been native to the southeast of the Arabian Peninsula. Archaeological data indicate that the dromedary was domesticated in the south of the peninsula in the second half of the second millennium BCE (decrease in bone size, increase of bones in settlement contexts, first undoubted figurines of dromedaries; Uerpmann & Uerpmann 2012; see also Heide 2011: 339–343, Almathen *et alii* 2016).² It is around the same time that the dromedary makes its first appearance in cuneiform texts: Nippur, Middle Babylonian, 14th–13th c. BCE **anše.a.ab.ba**, Ugarit ca. 1200 BCE [**anše.a.a]b.ba** (Heide 2011: 352, 368). The Sumerogram translates literally as ‘donkey of the sea’, a neologism which presumably indicates that it arrived in Mesopotamia by trade ship, hence in domesticated form.

² Potts (2004: 151 and 155) mentions dromedary remains at Tepe Ghabristan, Iran, during its Period 4 (3700–3000 BCE), which would be exceptionally early, as well as dromedary remains at Harappan sites of the later third and early second millennium BCE. It remains to be established whether the identifications and dates are reliable and if so, how they relate to data and ideas about the much later domestication of the dromedary in its native area in the south of the Arabian Peninsula.

In short, the Bactrian camel of the Central Asian steppes was domesticated perhaps as early as the late seventh millennium BCE in Inner Mongolia. The first definite traces of domesticated camels in southern Central Asia date from around the middle of the third millennium BCE, by which time their use spread rapidly towards the western steppe north of the Caucasus and via Iran towards the Middle East. This is at least a millennium earlier than the earliest secure evidence for the domestication of the dromedary, which was native to the southeast of the Arabian Peninsula. The literary evidence indicates that the Bactrian camel reached Mesopotamia by the middle of the third millennium BCE at the latest, while the dromedary made its first attested appearance over a millennium later.

4 Lexemes travelling with their denotata

This information is relevant to the theoretical places of origin of words for ‘camel’ and the routes that they may have travelled. Roughly two scenarios present themselves, one starting from the dromedary and the other from the Bactrian camel.

Scenario 1. The Semitic languages of the Arabian Peninsula may have had a very old word for the native dromedary that people encountered in the wild and no doubt hunted. This word, which under scenario 1 was the ancestor of the family of Hebrew *gāmāl*, spread over a wider area when dromedaries were domesticated and traded after approximately 1500 BCE. The word was subsequently also applied to the Bactrian camel, which had been known in southern Central Asia, in Mesopotamia and further south as a domesticate from the middle of the third millennium onwards. During the first millennium BCE the word was borrowed from Semitic into Greek, perhaps via an Anatolian language, and thence into Latin. This is approximately the scenario that agrees with the traditional narrative about the origin and spread of the word ‘camel’.

This scenario may perhaps be both supported and complicated by the fact that North African Berber and Tuareg possess a word for ‘camel’ that is very similar to **gam(a)l-* but with all consonants in a different position: this is Berber *aly^(w)əm*, Niger Tuareg *aylam*, which is discussed extensively by Kossmann 2005: 32–41 (he reconstructs **aa-IVqum*, where **aa-* is the obligatory nominal prefix).³

3 I am indebted to Maarten Kossmann for pointing me to the Berber forms and for discussion.

Scenario 2. An unknown language that was spoken somewhere between the eastern Central Asian steppes, where wild Bactrian camels lived and early domestication probably occurred for the first time, and Turkmenistan and eastern Iran, where domesticated camels appear in the third millennium BCE, had a word related with ‘camel’ (through inheritance or borrowing). This became widespread when domestication and long-distance transport made the Bactrian camel an economically important commodity. The word ended up in the Semitic world together with the animal itself, perhaps as early as the later third millennium BC, where it spread to become the word for ‘camel’ as well as ‘dromedary’ in the form of Hebrew *gāmāl* and its ilk. In that case *gāmāl* etc. replaced an earlier native word for (wild) ‘dromedary’.⁴ Since no attestation of Semitic *gāmāl* etc. is earlier than the first millennium BC (Militarev & Kogan 2005: 116–118), so more than 1500 years after the first attested appearance of domesticated Bactrian camels in Mesopotamia, this provides enough time for the scenario to unfold, rendering scenario 2 a theoretically conceivable scenario. Whether it is also a plausible scenario is a different matter: until cognates of the word family ‘camel’ surface on the Asian steppe and in the area north of Mesopotamia scenario 2 remains purely hypothetical.

5 ‘Camel’ in Caucasian languages

This is where the evidence of the Caucasian languages comes in. The Caucasian languages comprise three language families: West Caucasian, with the languages Abkhaz, Circassian and extinct Ubykh; South Caucasian, or Kartvelian, to which four languages belong, of which Georgian is the most widespread and has the largest number of speakers; and East Caucasian, which consists of around 30 languages in two major branches: the western branch is Nakh, which comprises Ingush, Chechen and Batsbi; all other languages belong to the eastern branch, Daghestanian, of which Avar-Andi-Tsezic forms a major group. It is in the Nakh and Tsezic languages of East Caucasian and in Georgian that we find hitherto insufficiently acknowledged relatives (by borrowing or inheritance) of the ‘camel’ word family.

⁴ A candidate for being an old word for ‘dromedary’ is the group of Sabean *?bl* (7th c. BCE, Arabian Peninsula) ‘dromedary’, Arabic *?ibil* ‘dromedary, Bactrian camel’, which is common in Semitic languages except for the Canaanite group (cf. Militarev and Kogan 2005: 4, Heide 2011: 346, 345 fn. 26).

In Caucasian studies, the existence of a word for ‘camel’ that is related to Semitic and Greek is recognized, for instance in Klimov & Xalilov 2003: 224. They regard Greek κάμηλος, which they state was itself borrowed from Semitic, as the direct source of Georgian *aklemi* ‘camel’, and Georgian *aklemi* as the source from which Chechen *emkal* and similar Nakh forms were borrowed. While it is reasonable to posit a relationship between those terms, given the overall similarity, the assumed route of borrowing gives rise to so many formal questions that a detailed investigation is called for. Let us first turn to the word for ‘camel’ in the Nakh languages, which has the most complicated history.

5.1 ‘Camel’ in Nakh

Standard Chechen, which is based on the dialects of the plains and foothills of the central North Caucasus, has nominative singular *emkal*, an oblique stem (from which all cases except the nominative are formed) *emkal(a)-*, and a nominative plural *emkal-š*.⁵ Similar forms are attested in the eastern Chechen Akki dialect (*emkal*, nominative plural *emkals̄*) and in the southern Chechen Kist dialect (*enkal*, nominative plural *enkališ*, which shows assimilation of *mk* to *nk* /ŋk/).⁶ In closely related Ingush, the word for ‘camel’ is *inkal*, oblique stem *inkal(a)-*, nominative plural *inkalaž*, which also shows assimilation of *mk* to *nk*. Batsbi has a word *aklam*, which will be addressed in section 5.3.

The history of the consonants in these words is relatively simple. Rather than the consonantal sequence *k-m-l*, which we encounter in the Semitic and Greek words for ‘camel’ that were discussed earlier, Chechen and Ingush show a sequence *m-k-l*. The consonants can be reconstructed as *m-k-l* for any stage between attested Nakh itself and Proto-East Caucasian (if the word existed here as early as that, which is of course not known, but see section 7; nor is the approximate date of Proto-East Caucasian known).

Reconstructing the vocalism is a different and more complicated story. First of all because Chechen and Ingush have undergone many vowel changes since the Proto-Nakh stage which come under the heading of umlaut (i.e. partial or total regressive vowel assimilation). Section 5.1.1. is devoted to the effects of umlaut on the word for ‘camel’ and to the reconstruction of Proto-Nakh. Deeper in time beyond that stage, the reconstruction of the vocalism becomes murkier but it needs to be addressed because the vowel changes that are known to have af-

⁵ Nichols and Vagapov 2004: 295.

⁶ Aliroev 1975: 89.

fectured Pre-Proto-Nakh are relevant to an assessment of the earliest shape of the word for ‘camel’. Those very early vowel changes are the subject of section 5.1.2.

5.1.1 Umlaut in Chechen and Ingush and the reconstruction of the Proto-Nakh word for ‘camel’

The rules of palatal and labial umlaut in Chechen and to a lesser extent Ingush have been worked out in the unsurpassed study by Imnajshvili 1977: 51–125. A number of precisions, especially those pertaining to the effects that the close vowels **i*, **u* had as opposed to the effects of the mid vowels **e* and **o*, were made by Schrijver 2021.

With few exceptions, Chechen and Ingush carry the stress on the first syllable. In many dialects, among those the dialects on which standard Chechen and Ingush are based, have lost all oppositions between old short vowels in unstressed, non-initial syllables, replacing them by *-a-* [ʌ]. This is the *-a-* in Chechen *emkal* and Ingush *inkal*, which consequently may go back to any of the Proto-Nakh short vowels **i*, **e*, **a*, **o*, **u*. We know that Proto-Nakh possessed those five vowels in non-initial syllables because the particularly archaic Cheberloj dialect of Chechen still preserves the opposition of all five short vowels, and so does the phonologically archaic Batsbi language.

The Chechen form *emkal* contains a clue to the original quality of *-a-*. In order to get a handle on this it is necessary to dwell on the phenomenon of palatal umlaut. The vowel *e*⁷ in initial syllables in Chechen always resulted from palatal umlaut of an old short **a*, which was caused by either **e* or **i* in the second syllable. So Chechen *emkal* reflects Proto-Nakh **amkel* or **amkil*. Which of the two is the correct reconstruction can normally be decided on the basis of the Ingush cognate because in Ingush **i* does (**a > e*) and **e* does not cause palatal umlaut of **a*.⁸ Like

7 This is *e* as opposed to *ie*, which in standard Chechen is never the result of umlaut. In standard Cyrillic orthography, *e* and *ie* are both spelled as <e>, which obscures this important opposition. The dictionary by Nichols and Vagapov (2004), the phonological description by Nichols 1997, the grammatical description by Nichols 1994, the Chechen morphology by Jakovlev (1960) and some other studies do distinguish *e* from *ie*. A similar situation obtains for Ingush, where *e* and *ie* are spelled in standard Cyrillic as <e>. In spoken Ingush, short **ie* (which usually but not exclusively occurs in closed syllables) merged with the *e* that arose as a result of umlaut (Nichols 2011: 25), which probably is a very recent phenomenon (Imnajshvili 1977 still distinguishes the two).

8 The different umlaut effects of **e* and **i* on **a* in Ingush are not specifically identified by Imnajshvili (1977), but they can be distilled from the Chechen material he provides if one compares Ingush cognates, e.g. Proto-Nakh **bali* ‘shoulder(s)’ > Chechen *bela-š*, Ingush *belaz*, both with umlaut (Cheberloj *bališ* and Batsbi *bali* preserve the original vocalism in first and second syllable); Proto-Nakh

standard Chechen, Ingush lost the opposition between the old short vowels in non-initial syllables, replacing them by *a* [ʌ]. Unfortunately, however, Ingush is not so obliging in this instance: its word for ‘camel’ is *inkal*, oblique stem *inkal(a)-*, nominative plural *inkalaž*, which shows *i-* rather than expected **e* (if from **amkil*) or **a* (if from **amkel*). The correspondence Ingush *i* ~ Chechen *e* is irregular and cannot be explained on the basis of known sound changes involving umlaut. For the time being, there is no other option than to reconstruct Ingush *inkal* as Proto-Nakh **imkVl*, with old **i* in the first syllable and with **V* in the second syllable representing any one of the Proto-Nakh vowels **a*, **e*, **o* or **i*.⁹ We shall return to Ingush in section 5.1.2. So we cannot decide whether Chechen *emkal* reflects **amkel* or **amkil*.

Beside **amke/il* (Chechen) and **imkVl* (Ingush), a third form, **amkal*, must have existed, as is revealed by Cheberloj Chechen *ankal* and Itumkali Chechen *ankal* (Aliroev 1975: 89). As stated earlier, Cheberloj preserves Proto-Nakh vowel quality oppositions in non-initial syllables, so the second *-a-* in *ankal* represents original *-a-*. Itumkali, on the other hand, lost non-initial short vowel oppositions and underwent palatal umlaut much like standard Chechen did. The fact that Itumkali *ankal* does not show the effect of umlaut indicates that underlying it is the same form as is attested in Cheberloj, viz. Proto-Nakh **amkal*.

So at the end of this discussion, we arrive at the necessity to reconstruct three different forms of the word for ‘camel’ for Proto-Nakh:

1. **amkel* or **amkil* (Plains Chechen)
2. **amkal* (Cheberloj and Itumkali Chechen)
3. **imkVl* (Ingush)

5.1.2 Pre-Nakh origins of the Nakh vowel alternation **i* ~ **a*

In the Proto-Nakh reconstructions we observe an alternation in the first (stressed) syllable between **i* and **a*. In the second syllable, **e* or **i* alternates with **a*. Neither alternation can be explained on the basis of umlaut, but they do have an equivalent in the nominal morphology of the Nakh languages. Consider the fol-

**malin* ‘warm’ > Chechen *mela^o*, Ingush *mela* (Cheberloj *mali^o*, Batsbi *mali^o*); contrast Proto-Nakh **maqe* ‘harrow’ > Chechen *meqa* (umlaut), Ingush *maqa* (no umlaut; Cheberloj *maqe*); Proto-Nakh **laqen* ‘high’ > Chechen *leqa^o*, Ingush *laqa* (Cheberloj *laqe^o*, Batsbi *laqe^o*). See Schrijver 2021: 95–100. Nikolayev and Starostin (1994: 99) record the difference but they do not provide supporting material. ⁹ Not **u* because this would have caused labial umlaut resulting in Ingush **unkal*, cf. =*ug* ‘leads’ < **=ik’-u* of the Proto-Nakh verbal root **=ik’*, which is still represented as such in Batsbi =*ik’-*).

lowing examples (abbreviations: E = ergative case, G = genitive case, Loc. = one of the local cases, O = oblique case stem):

Proto-Nakh	Batsbi	Chechen	Ingush	meaning
1. *dik' O *dak'or-	dik' pl. dak'vri' beside dik'ujr	dig G dagara ^a E dagaruo	dig E dogaruo	'axe'
The Batsbi plural forms are those of Desheriev 1953: 68 (Kadagidze & Kadagidze 1984 only provides the nominative singular); the nominative plural <i>dak'vri'</i> probably represents * <i>dak'ujri'</i> < * <i>dak'or-i</i> , with regular umlaut according to Batsbi rules (cf. Imnajshvili 1977: 120); <i>dik'ujr</i> reflects the generalization of the first-syllable vocalism of the nominative singular <i>dik'</i> . The Chechen and Ingush forms are those of Nichols & Vagapov 2004 and Nichols 2004, respectively. The Proto-Nakh reconstructions take the effects of umlaut in Chechen and Ingush into account. For the etymology see Giginejshvili 1977: 82 (on Lak and Lezgian cognates), Nikolayev and Starostin 1994: 944, Nichols 2003: 258.				
2. *niq O *naqorV-	-	niq G naqara ^a E naqaruo	niq E noqaruo	'beehive'
The Chechen and Ingush forms are those of Nichols & Vagapov 2004 and Nichols 2004, respectively. See Nikolayev & Starostin 1994: 868 for the etymology.				

In both examples, **i* is found in the first syllable of the nominative singular, while the oblique stem, from which cases like the genitive, ergative and dative are formed, and sometimes also the plural, instead has **a*. The **a* is preserved without change in Batsbi and in Chechen, but Ingush has turned it into *o* as a result of labial umlaut caused by original **o* in the second syllable of the oblique stem (in standard Chechen as well as in the Plains dialect, **o* does not cause labial umlaut of **a* in the preceding syllable, cf. Imnajshvili 1977: 65–68, Schrijver 2021: 108–110). Subsequently, the unstressed **o* regularly became *a* [ʌ] in Ingush and Chechen.

This alternation of **i* in the nominative singular stem and **a* in the oblique stem is relatively rare in nominal paradigms, but it has counterparts which are more frequent and in which the nominative singular contains **o* or **u* and the oblique stem is characterized by **a*, e.g.¹⁰

¹⁰ Batsbi forms are taken from Kadagidze and Kadagidze 1984, Gagua 1961: 85, Holisky-Gagua 1994: 161, 167. Chechen forms from Nichols and Vagapov 2004 and Ingush forms from Nichols 2004.

Proto-Nakh	Batsbi	Chechen	Ingush ¹¹	meaning
3. * <i>borc</i> O * <i>barci-</i>	<i>borc</i> G <i>barci</i> ⁿ E <i>barcav</i>	<i>buorc</i> G <i>berca</i> ⁿ E <i>bercuo</i>	<i>buorc</i> E <i>bercuo</i>	'millet'
4. * <i>mot't</i> G * <i>mat't'i</i> ⁿ O * <i>mat't'a-</i>	<i>mot't</i> G <i>mat't'i</i> ⁿ E <i>mat't'av</i>	<i>muott</i> G <i>metta</i> ⁿ E <i>mattuo</i>	<i>muott</i> E <i>mettuo</i> ^a	'tongue'
5. * <i>buc</i> G * <i>ba:ci</i> ⁿ O * <i>ba:ca-</i>	<i>buc</i> G <i>baci</i> ⁿ Loc. <i>bac-</i> <i>ma-k</i>	<i>buc</i> G <i>be:ca</i> ⁿ E <i>ba:cuo</i>	<i>buc</i> E <i>beacu</i> ^a	'grass'
6. * <i>butt</i> G * <i>batti</i> ⁿ O * <i>batta-</i>	<i>butt</i> G <i>batti</i> ⁿ E <i>battav</i>	<i>butt</i> G <i>betta</i> ⁿ E <i>battuo</i>	<i>butt</i> E <i>bettuo</i> ^a	'moon'

For reasons that remain to be worked out, the *a* in such paradigms appears as long *a*: in some etyma, e.g. in **buc* oblique **ba:c-* 'grass'. As argued by Schrijver 2021: 131–146, the Nakh vowel alternation of **i*, **u*, **o* in the first syllable of the nominative singular with **a* (or **a:*) in the first syllable of the oblique cases and the nominative plural has deep roots, which go back all the way to Proto-East-Caucasian, the common ancestor of Nakh and the Daghestanian languages. Nakh has a particularly close counterpart in the Tsezic languages of Daghestanian.

Proto-Nakh	Batsbi	Chechen	Ingush	Tsezic
7. * <i>mott</i> O * <i>matt-</i> 'bed, place'	<i>mott</i> Adessive <i>matteḥ</i> pl. <i>mattiš</i>	<i>muott</i> G <i>metta</i> ⁿ E <i>mattuo</i>	<i>muott</i> E <i>mettuo</i>	* <i>mɔč:ʷ</i> O * <i>mič:ʷ-</i> > * <i>mič:ʷ-</i> in Tsez <i>mōči</i> O <i>mečo-</i> , <i>moči-</i> , Hunzib <i>mɔče</i> O <i>mičo-</i> 'place, plot'
Etymology: Nikolayev & Starostin 1994: 803.				
8. * <i>duq</i> ' pl. * <i>daq'w-</i> 'yoke; mountain crest'	<i>duq</i> ' 'yoke'	<i>duq</i> ' G <i>duq'a</i> ⁿ E <i>duq'uo</i> pl. <i>daq'q'aš</i>	<i>duq</i> ' E <i>duq'uo</i> pl. <i>douq'až</i>	* <i>ruλ'u</i> > Bezhta <i>ruλ'o</i> ; * <i>riλ'we-</i> in Tsez <i>ra'λ'u</i> , <i>roλ'i</i> , Hunzib <i>riλ'u</i> 'yoke'
Etymology: Giginejshvili 1977: 109, Nikolayev & Starostin 1994: 220, 954, Nichols 2003: 260.				

Where Nakh has an alternation nominative **u*, **o* ~ oblique stem **a*, the Tsezic languages show an alternation **u*, **o*, **ɔ* ~ **i* in cognate etyma. On the basis of a close comparison of the Tsezic data with those from Nakh it is possible to establish that Proto-East Caucasian had the following features (Schrijver 2018a on Avar-Andi-Tsezic, Schrijver 2021 on Nakh and Proto-East Caucasian):

¹¹ Palatal umlaut in the Ingush ergative is probably due to the generalization of the stem-final vocalism of the genitive.

- (1) it possessed phonological stress, much like present-day Avar;
- (2) nominal paradigms existed in which the nominative singular had the stress on the first syllable while the oblique stem had stress on a non-initial syllable (this is a type commonly attested in Avar, e.g. *moč*:’, genitive singular *moč:’ról* ‘moon’);
- (3) the rounded vowels **u*, **o*, **ɔ* were regularly unrounded to **i* in pretonic position in Proto-East Caucasian; hence the Proto-East Caucasian nominative (= absolutive) preserved the rounded vowel because it was stressed, while the oblique stem turned it into **i* because it was pretonic.

With this original state of affairs in mind, the vowel alternation in Nakh receives a relatively simple explanation on the basis of the following regular vowel correspondences and reconstructions (Schrijver 2021: 138–149):

Proto-East Caucasian		<i>*u</i>	<i>*o</i>	<i>*ɔ</i>	<i>*i</i>
Proto-Nakh	stressed and posttonic	<i>*u > *u</i>	<i>*o > *u</i>	<i>*ɔ > *o</i>	<i>*i > *i</i>
	pretonic	<i>*u > *i > *a</i>	<i>*o > *i > *a</i>	<i>*ɔ > *i > a</i>	<i>*i > *a</i>

It is the last column in this diagram that is relevant to understanding the Proto-Nakh paradigms of **dik*’, oblique **dak’or-* ‘axe’ and **niq*, **naqor-*, which resulted from Pre-Proto-Nakh **dik*’, oblique **di’k’or-* and **niq*, oblique **ni’qor-*, depending on the original position of the stress.

In a few examples, the vowel alternation **i ~ *a* is not only attested in the first syllable but also in the second, which suggests a more complex form of accentual mobility according to which the stress moved between the first, second and third syllables and consequently the second syllable was posttonic in the nominative, stressed in some oblique forms, and pretonic in others:

Proto-Nakh	Batsbi	Chechen	Ingush	Pre-Proto-Nakh
9. <i>*jiš</i> O <i>*a:šir-</i> , <i>*a:šar-</i> ‘voice’	<i>iš</i> E <i>aširv</i>	<i>jiš</i> G <i>e:šara</i> ^m E <i>e:šaruo</i>	<i>jiš</i> E <i>a:šaruo</i>	<i>*iš</i> O <i>*i’širV-</i> , <i>*iš’rV-</i>

Batsbi shows *i*-vocalism in the second syllable of the oblique stem, which agrees with palatal umlaut in the Chechen oblique stem. Ingush lacks palatal umlaut, which prohibits the reconstruction of **i* in the second syllable and necessitates the reconstruction of **a* (this is the only vowel which in Ingush does not cause umlaut of **a*: in the preceding syllable, cf. Schrijver 2021: 118). In accordance with the stress-dependent behaviour of Proto-East Caucasian **i* in Nakh, the deeper reconstruction of **jiš*, O **a:šir-*, **a:šar-* was probably **iš* O **i’širV-*, **iš’rV-*, respectively, which yielded Proto-Nakh **iš* O **a:širV-*, **a:šarV-*.

(continued)

Proto-Nakh	Batsbi	Chechen	Ingush	Pre-Proto-Nakh
10. *jis O *a:sir-, *a:sar- 'hoarfrost'	–	jis G e:sara ⁿ E e:saruo	jis E a:saruo	* ¹ is O * ¹ ʃsirV-, * ¹ is ^ʃ rV-
The situation is similar to that of 'voice' except that <i>i</i> -vocalism of the second syllable is not directly attested in Batsbi. For the etymology outside Nakh see Nikolayev & Starostin 1994: 675.				
11. *ʃi O *ʃanar-, *ʃanir- 'steam'	ʃa, Lative ʃanar- <i>ɬ</i>	ʃa G ʃänara ⁿ	ʃi E ʃanaruo	* ¹ ʃi O * ¹ ʃi ⁿ rV-, * ¹ ʃi ⁿ rV-
The Nakh oblique stem *ʃanar- is required for Batsbi, while *ʃanir- is presupposed by Chechen, which shows palatal umlaut (<i>ä</i> instead of <i>e</i> is regular in pharyngeal environment, see Imnajshvili 1977: 60, Schrijver 2021: 95–100). The absence of umlaut in Ingush ʃanar- suggests that it generalized the stem *ʃanar-. Ingush is the only language to preserve the root vowel <i>i</i> in the nominative. See Nikolayev & Starostin 1994: 485 for possible etymological connection with Daghestanian.				

Consequently, what these paradigms show is the stress-related behaviour of Proto-East Caucasian **i* in Nakh: it became Proto-Nakh **i* under the (Proto-East Caucasian) stress and **a* in pretonic position.

As a final twist, it seems that we are not justified in automatically reconstructing Proto-East Caucasian **i* wherever the Nakh alternation **i* ~ **a* occurs, as I have done in the diagrams. The reason for that is that in the case of two etymologies involving Nakh **i* ~ **a* the evidence from Avar-Andi-Tsezic suggests the reconstruction of Proto-East Caucasian **a* rather than **i*:

- Proto-Nakh *jis O *a:sir-, *a:sar- 'hoarfrost' is related to Andi *asor*, Northern Axwax *aša* < Proto-Andic **asor*. Andi *a* can only reflect earlier **a*. The same vocalism is found in further cognates: Avar *fansi* 'snow drift', Tsez *az-q'a* (with unclear *-q'a*), Hunzib *a^oza* < **ansV*. See Nikolayev & Starostin 1994: 675 for the etymology.
- Proto-Nakh *ʃi O *ʃanar-, *ʃanir- 'steam' is probably related to Andi *hal* 'steam', which has **a*. A further relative is the verbal stem **hal-* (Avar *hal-*, *hwal-* 'boil'), **hel-* (Bezhta and Hunzib *hel-* 'to boil'). See Nikolayev and Starostin 1994: 485.

So in as far as can be judged on the basis of just two etymologies, it seems that nouns that had original **a* in the first syllable could join the alternation of nominative **i* ~ oblique **a* in Nakh. It is unclear whether this was a result of regular sound change (i.e. Proto-East Caucasian stressed **a* becoming **i* > **i* in Nakh) or rather of analogy (i.e. the morphological pattern nominative **i* ~ oblique **a* (< **i*) becoming productive at the expense of the pattern nominative **a* ~ oblique **a* (< **a*)).

On the basis of this information, it is possible to propose that the different reconstructions of the word for 'camel' in Nakh that we found in section 5.1.1 orig-

inally formed a single paradigm that looked like the paradigms of the words for ‘voice’, ‘steam’ and ‘hoarfrost’:

- **imkVI* (Ingush) can be reconstructed as the Proto-Nakh nominative, **imkil*
- **amkel* or **amkil* (Plains Chechen) can be reconstructed as one of the two oblique stems, **am'kil-*
- **amkal* (Cheberloj and Itumkali Chechen) can be reconstructed as the other oblique stem, **amka'IV-*

The original paradigm **imkil*, oblique **amkil-* beside **amkaIV-* was apparently so irregular that the various dialects of Checheno-Ingush generalized a single stem throughout the paradigm, but they did so by selecting a different stem.

5.1.3 A formal parallel to ‘camel’ in East Caucasian

There exists another etymon which in the Nakh languages shows an alternation between **i-* in Ingush and **a-* in Chechen, thus offering a close formal parallel to ‘camel’. I am referring to the word for ‘glass, porcelain’:¹²

- Chechen *angali* (Matsiev 1961), *āngli*, *ā:ngli* (Nichols-Vagapov 2004) ‘glass (shard), porcelain’. The Chechen dialectal forms Kist *engli*, Cheberloj *angali* and Itumkali *āngle* are provided by Aliroev 1975: 41. These forms reflect **a(:)ngalaj*.
- Ingush *ingalii* (Nichols 2004) ‘porcelain’, from **ingalaj*

The interesting aspect of this etymon is that it has Daghestanian cognates which directly point to a paradigmatic alternation between Proto-East Caucasian **i* and **ɔ* and hence provide evidence for the idea that the Nakh alternation **i ~ *a* indeed stems from Proto-East Caucasian **i*: for Avar-Andi-Tsezic it is possible to reconstruct a paradigm absolutive **hɔŋV(r)*, oblique stem **hɪŋwVr(V)-*.¹³

- The vocalism of the first syllable of the absolutive was generalized in the Tsezic languages: Tsez and Hinuq *aki*, oblique stem *aki-* ‘window’; Xwarshi *aⁿka* (Imnajshvili 1963: 16), Inxoqwar *aⁿka*, *aka*, oblique stem *aⁿka-* ‘window, opening, hole’ (Khalilova 2009: 89, 90; Kibrik & Kodzasov 1990, item 375). These forms reflect Proto-Tsezic **ankV(r)* from earlier **hɔŋV(r)*, with **a < *ɔ* which is regular in word-initial position before nasal + plosive (/#_NT).¹⁴

¹² I am indebted to an anonymous referee for pointing me to this etymon.

¹³ The reconstruction of **h-* rather than another pharyngeal or laryngeal consonant is arbitrary.

¹⁴ Direct evidence for a Tsezic rule **ɔ > *a* /#_NT is provided by (1) Tsez *aqju*, Inxoqwar *aⁿqqu*, Hunzib *aqqu* ‘female (animal)’ < **anq-ju*, Hinuq *aqili*, Xwarshi *aqu*, Bezhta *aqo*, Hunzib *aqe* ‘woman’ < **anqə*; cf. Andi *horč:i*, Botlix *hark:i* ‘wife’ < **hɔrk:i*; (2) Tsez, Hinuq *ata*, Xwarshi

- The Andic languages generalized the vocalism of the first syllable of the oblique stem: Andic *ingur* (Salimov 2010: 300), Northern Axwax *íngo*, Southern Axwax (Tsegob, Tljanub) *ingo*, Southern Axwax (Ratlub) *ungor* ‘window’; Chamalal *inna* ‘opening, hole, window’, Chamalal (Gigatl) *hüngár* ‘window’, Tindi *hìngwar* ‘window, kind of ring-shaped bread roll’, Karata *hüngwár*, Karata (Tokita) *hingur* ‘window; kind of pasty’, Botlix *híngur*, Bagwalal (Xuštadi) *húngar*, Bagwalal (Tlis-sib) *húngwa* ‘window’, Godoberi *hingúr* ‘window, pasty’ (e.g. Kibrik & Kodzasov 1990, item 375, and the many dictionaries that appeared since the 1990s).

Further Daghestanian cognates: Lezgi *hamga* ‘crystal’, Tabassaran *hamg*, Aghul *ŕagug* ‘glass’, Kryz *ŕag* ‘niche in wall, roof orifice’, Budux *ħag* ‘roof orifice’ (cf. Nikolayev & Starostin 1994: 536).

The etymon does face the problem that the Batsbi cognate, *ank’raʔǝ* ‘glass’ (Kadagidze & Kadagidze 1984: 40), has *k’* rather than *g*, which would indicate that Chechen and Ingush *g* reflects Proto-Nakh **k’* rather than **g*, thus providing an obstacle to the Nakh-Daghestanian equation. However, the Batsbi word also means both ‘pure, clear’ and ‘grass-snake’. The similarity of these Batsbi forms to Georgian *ank’ara* ‘clear, pure (of water, voice)’ and ‘grass-snake’ must be more than accidental, but the exact nature of the relationship is unclear: perhaps Georgian donated the item to Batsbi, or the native Nakh etymon for ‘glass’, which had **g*, was formally influenced by its Georgian look-alike and replaced **g* by **k’*.

5.2 ‘Camel’ in Georgian

Before delving deeper into the history of the Nakh word for ‘camel’ it is necessary to study the other Caucasian languages that have cognates.

The Georgian word for ‘camel’ is *aklemi*. This has the same initial vowel *a-* as two of the Nakh forms but a different vowel in the second syllable (*e* rather than Nakh **a* and **i*). It also differs from Nakh by showing a different ordering of its consonants: *k-l-m* rather than Nakh *m-k-l*, an issue to which we shall return in section 5.5. *Aklemi* has an inflectional stem *aklem-*, the final *-i* being the productive suffix that was added to the Nsg. of all consonantal stems.

a^sta, Inxoqwar *a^sta*, *a^sta*, Bezhta *ä^sdä*, Hunzib *a^sda* ‘brain’ < **a^snda*, cf. Andi *honu* ‘brain’, Botlix *handa* ‘forehead’ < **hɔn(d)V*. Indirect evidence is provided by numerous instances in which Tsezic **a* /#_NT corresponds to Andic **i* (with **i* < Proto-East Caucasian unrounded pretonic **ɔ*), e.g. Tsez, Hinuq *ac*, Xwarshi, Inxoqwar, Bezhta *a^sc* ‘door’, Andi, Botlix *hinc:u* ‘id.’; Tsez, Hinuq *a^stu*, Inxoqwar *a^ste*, Bezhta *a^sto* ‘armful’, Andi *intir*, Botlix *intar* ‘bosom’.

Closely related Mingrelian has *arkemi* ‘camel’, which Klimov & Xalilov 2003: 224 regard as a borrowing from Georgian.

5.3 ‘Camel’ in Bezhta and Batsbi

Georgian *aklemi* was probably borrowed in two East Caucasian languages. One is Bezhta, a Daghestanian language that belongs to the Tsezic group of the Avar-Andi-Tsezic subfamily. This has *aklamo* ‘camel’ (Xalilov 1995: 26). Speakers of Bezhta are in frequent contact with Georgian speakers and the language contains many Georgian loanwords. It is therefore reasonable to suppose, with Komri, Xalilov & Xalilova (2015: 77), that Bezhta borrowed the word for camel from Georgian. They provide more examples in which Georgian *e* was borrowed as Bezhta *a*, e.g. Georgian *sabeli* ‘rope’ > Bezhta *sabali*; all instances have an *a* in the first and *e* in the second syllable, which is assimilated to *a* - *a* in Bezhta. The final *-o* is an innovation of Bezhta, the motivation for which is unclear.

The other language that probably borrowed from Georgian is the Nakh language Batsbi, which has *aklam* ‘camel’ (Kadagidze & Kadagidze 1984: 50). Speakers of Batsbi are all fluent in Georgian and the language abounds with Georgian loanwords. Such loanwords regularly lose the final *-i* of the nominative of Georgian consonantal stems, but there is no apparent reason why *aklem-* should have become Batsbi *aklam*. In general, Batsbi has many instances of the sequence *a - e*, e.g. *bader* ‘child’, *pɣaner* ‘shoulder’, *sak'er* ‘neck’. It is perhaps conceivable that *aklam* represents a cross between borrowed *aklem-* and the native Nakh stem *amkal-*.

5.4 ‘Camel’ in Tsez and Hinuq

The final representatives of ‘camel’ in Caucasian languages are found in two other Tsezic languages: Tsez *ɣumukúli*, the Asax dialect of Tsez *ɣomokuli*, and Hinuq *ɣomókilu* ‘camel’. These forms share the consonantal sequence *m-k-l* with the Nakh languages and the fact that Tsez and Hinuq are spoken in the southwestern part of Daghestan that is close to Chechnya suggests that they may have been borrowed from Chechen at some prehistoric stage. The quadrisyllabic stem structure is highly unusual in Tsezic languages (one or two syllables are the norm). The similarity of the initial syllables to the West-Tsezic word for ‘donkey’, Tsez *ɣomɔj*, Hinuq *ɣomoq'i*, Xwarshi *umuq'e*, Inxoqwar *om'oq"e* < Proto-Tsezic **ɣomɔq'e* suggests that the loanword for ‘camel’ was remodeled by popular etymology on the native word for ‘donkey’ (for a semantic parallel for the association of camel and donkey, cf. Middle Babylonian **anše.a.ab.ba** ‘dromedary’, literally ‘donkey of

the sea', see section 3). The final syllables in Tsez, *-kuli*, do not match those of Hinuq, *-kilu*. One approach to this difference is to assume metathesis of the vowels in one of the languages, but this has the drawback of being *ad hoc*. Alternatively, a common Tsez-Hinuq stage **-kwil-* may be reconstructed, which regularly became *-kul-* in Tsez and *-kwil-* in Hinuq, as in the following native etyma:¹⁵

Tsez	Hinuq	Proto-Tsez-Hinuq	Proto-Tsezic	meaning
<i>Ɂut-</i>	<i>Ɂwit-</i>	<i>*Ɂwit-</i>	<i>*Ɂwet-</i>	'to perish, die'
<i>=iħu-</i>	<i>=iħwi-</i>	<i>*=iħwi-</i>	<i>*=eħ:we-</i>	'to collect (in a stable), sit (on eggs)'
<i>=a^ɕqu-</i>	<i>=a^ɕwi-</i>	<i>*=a^ɕwi-</i>	<i>=a^ɕ:we-</i>	'to sew'
<i>q'u'l</i>	<i>q'wil</i>	<i>*q^ɕwil</i>	<i>*q^ɕwel</i>	'rind, bark'
<i>kud</i>	<i>kwid, kud</i>	<i>*kwid</i>	<i>*kwed</i>	'basket'

The problem facing this reconstruction is that Hinuq and the Asax dialect of Tsez normally retain **wi* unchanged (Imnajshvili 1963: 28), but it may well be that Asax *šomokuli* was borrowed from another Tsez dialect. The difference in the word-final vowels between Tsez (*-i*) and Hinuq (*-u*) may reflect a different incorporation of this foreign lexeme into native inflexional morphology.

5.5 Conclusions on the attested Caucasian forms

It is possible to identify two related groups of words for 'camel' in Caucasian languages. The first has a consonantal sequence *k-l-m* and centres on Georgian *aklemi*. This, or a form very much like it, was probably borrowed into the East Caucasian languages Bezhta (*aklamo*) and Batsbi (*aklam*). The other group has a consonantal sequence *m-k-l* and centres on Checheno-Ingush. There we find a number of representatives whose vowel alternations are reminiscent of the vowel alternation in paradigms of native words such as Proto-Nakh **iš O *a:širV-*, **a:šarV-*, allowing the reconstruction of a Proto-Nakh paradigm **imkil*, oblique **amkilV-*, **amkalV-*. This means that if the word for 'camel' is a loanword in Nakh itself, it was well integrated into the morphology of the receiving language. The Nakh form **amkil-* may be the source for the borrowing into Tsez and Hinuq, which can be tentatively reconstructed as **šomokwil-*. Formal differences between **amkil-* and **šomokwil-* can at least partly be explained if we assume that the Tsez-Hinuq form was influenced by the word for 'donkey', **šomoq'e*.

¹⁵ Nikolayev and Starostin 1994 mention three of the etymologies (pp. 638, 701, 931). All items were checked against Xalilov 1999 and Xalilov and Isakov 2005.

The relationship between Georgian *aklemi* and Nakh **imkil*, oblique **amkil*, **amkalV*- is more difficult to establish. It makes little sense that Georgian borrowed from Nakh or Nakh borrowed from Georgian: there are no phonotactic reasons in either language why the **m* should move from the left-hand side to the right-hand side, or vice versa, skipping across the medial *-l*. It would seem that the only reason for assuming that the direction of borrowing was from Georgian to Nakh (Klimov-Xalilov 2003: 224) is the assumption that the word originated in Semitic and Georgian geographically lies between Semitic and Nakh. But it is not self-evident that the word originated in Semitic (see section 4), nor how the formal transformations between Semitic *gamal* and the like on the one hand and Georgian *aklemi*, Nakh **imkil*, oblique **amkil*, **amkalV*- on the other should be explained.

What is plausible is that the missing link between Georgian *aklemi* and Nakh *amkil*- etc. is a form with the *-m*- between *-k*- and *-l*-, so exactly where it is in Semitic *gamal* and Greek κάμηλος. It stands to reason that if a borrowing of the shape **akmil*- or the like entered Nakh, it would be adapted as **amkil*-: consonant clusters of the type plosive + resonant are not tolerated (unless there is a morpheme boundary in between), while clusters of the type resonant + plosive are very common (see Nichols 2011: 89–98 on such consonant clusters in Ingush; she does not mention clusters of nasal resonant + plosive, but they are not uncommon, e.g. *q'oandz* ‘vinegar’, *c'andar* ‘clean’, *c'onkar* ‘mace’, *enži*: ‘coarse’, *engar* ‘co-wife’, *onk* ‘good quality’, *sonka* ‘mountain crest’, all from Nichols 2004). It is less clear, however, why in Georgian a borrowing of the shape **akmel*- should become **aklem*- except by a sporadic metathesis. However that may be, as far as the consonants are concerned, an originally medial position of the *m* offers the best opportunity to account for its position in Georgian and Nakh.

How about the vocalism? Is it possible to propose a form that is capable of bringing together Georgian *aklemi* and Nakh **imkil*, oblique **amkil*-, **amkalV*-? Recall that if the Nakh item were native, its pre-Proto-Nakh reconstruction would be **imkil*-, with the development of **i* to Proto-Nakh **i* or **a* depending on stress, which was mobile in this paradigm. Also recall that nouns with an original **a* rather than **i* in the first syllable probably joined this pattern of noun inflection, so that beside a reconstruction **imkil* also **amkil* is possible. It is either the latter, Pre-Proto-Nakh **amkil*-, or the later allomorph, Proto-Nakh **amkil*-, that comes closest to Georgian *aklem*- in as far as the vowels are concerned.

In conclusion, the original form of the word for ‘camel’ that is best able to account for the reflexes in Nakh and in Georgian is either **akmil*- or **amkil*-, with *-m*- as the middle consonant and *a* and *i/i* as the vowels of the first and second syllable respectively.

6 The word ‘camel’ and a European lexical substrate

The question remains in which relationship Caucasian **akmil-*, which I shall henceforth use as a shorthand summarizing the conclusion of section 5 on Caucasian, stands to Semitic **gam(a)l-* and Greek κάμηλος. This brings us to the heart of the reason why this study on the history of the word ‘camel’ forms part of a collection of essays on the languages of pre-Indo-European Europe. For the formal relationship of **akmil-* to **gam(a)l-* and κάμηλος is strongly reminiscent of a set of etymologies that are believed to be of pre-Indo-European origin and that are widespread throughout Europe (Iversen & Kroonen 2017, Schrijver 2018b: 362–363, with references). I am referring here to a specific feature of some of those etymologies, viz. the presence or absence of an initial **a-*, which, if present, usually produces vowel changes or vowel loss in the body of the word. Here are a few relevant examples:

with <i>*a-</i>	without <i>*a-</i>
1. Old High German <i>amsala</i> , <i>amusla</i> ‘blackbird’ < <i>*amsl-</i>	Latin <i>merula</i> , Welsh <i>mwyalch</i> < <i>*mesal-</i> ‘blackbird’
2. Greek ἀστραπή ‘lightning’ < <i>*astrp-</i>	Cypriote Greek στροπά ‘id.’, Old Irish <i>sraif</i> ‘sulphur’ < <i>*str(a)P-</i>
3. Old High German <i>aruz</i> ‘ore’ < <i>*arud-</i>	Latin <i>raudus</i> ‘ore’ < <i>*raud-</i>
4. Welsh <i>erfin</i> ‘turnip’ < <i>*arb^h-</i>	Old High German <i>ruoba</i> , Latin <i>rāpa</i> , Greek ράφανος ‘turnip’ < <i>*rāp-</i> , <i>*rab^h-</i>
5. Greek ἀχράς, -άδ- ‘wild pear’ < <i>*a^grd-</i>	Albanian <i>dardhë</i> ‘pear’ < <i>*^ga/ord-</i>
6. Hittite <i>alel</i> ‘flower’ < <i>*aleil-</i> , Old Irish <i>ailestar</i> ‘flag-iris’ < <i>*alil-stro-</i>	Greek λείριον, Latin <i>lilium</i> ‘lily’ < <i>*leil-</i>

To this we can now add:

7. Caucasian <i>*akmil-</i> ‘camel’	Greek κάμηλος, Semitic <i>*gam(a)l-</i> ‘camel’
-------------------------------------	---

For some of those etyma Near Eastern counterparts have already been identified. The word for ‘ore’ (3) closely resembles Sumerian **urudu** /aruta/ (Schrijver 2018b: 363). The Hittite word for ‘flower’ (6) is usually connected to Middle Egyptian *hrr-t*, Coptic *hrēri*, *hlēli* ‘flower’, perhaps Berber *alili* ‘oleander’ (e.g. Tischler 1977:16, Puhvel 1984: 32–33). Kroonen 2012 drew attention to Greek γέλις, ἄλις ‘garlic’, possibly from **gedl-*, **a-gdl-* respectively and its relationship to Old Babylonian

gidlu ‘plaited string of onions and garlic’. So the word for ‘camel’ is not unique in this respect.

The linguistic stratum to which this group of originally non-Indo-European words belongs has been tentatively connected with the language or languages of the first farmers who entered Europe from Asia Minor from the seventh millennium BCE onwards (Iversen-Kroonen 2017, Schrijver 2018b). The fact that related lexemes are widespread in the Near East chimes well with this idea. The word for ‘camel’ obviously would not belong to this ancient stratum: on present evidence, the presence of domesticated Bactrian camels in Turkmenistan and subsequently on the western steppe and in the Near East cannot be pushed back beyond the third millennium BCE. But given its strikingly similar phonological pattern, it is possible that ‘camel’ originated in or, if it was a loanword, was morphophonologically adapted to the same language family to which the early European agriculturalists belonged and which may consequently be situated somewhere in the larger Near East (i.e. including Turkey, Iran, the Caucasus, Mesopotamia and the Levant).

Another interesting aspect of the word for ‘camel’ is that Caucasian languages are involved. It is quite possible that the word was borrowed there, but on the other hand it is striking how well the pattern found in the European substratal words with and without **a-* fits in with the morphophonology of reconstructed East Caucasian. Although the systematic reconstruction of that family is still in its infancy compared to the reconstructions of Indo-European, Uralic or Semitic, it is possible to outline features that were typical of the inflection of nouns. Of particular relevance is the behaviour of nouns of the original structure **(C)V₁CV₂C*. Those nouns lose either the entire initial syllable (especially if that syllable has a weak consonantal onset, i.e. a laryngeal or pharyngeal), or they lose the vowel of the second syllable. The conditions under which either occurs remain to be worked out, but it is probable that some case forms lost the first and others the second syllable, no doubt in connection with accentual mobility. Consider the following examples of etyma which show both developments:

<i>*(c)vqvc-</i>	<i>*(c)vqv/c-</i>	<i>*(c)vqv/c-</i>	Nikolayev & Starostin 1994
1. <i>*vmvrg-</i> ‘bracelet’	Tsez <i>aka</i> , Hunzib <i>a^hga</i> < <i>*amrg-</i>	Avar <i>mergó</i> , Godoberi <i>mingwa</i> , <i>Axwax magwa</i> < <i>*mvrg-</i>	NS 234
2. <i>*hvnvc:’-</i> ‘door’	Tsez <i>ac</i> , Bezhta <i>a^hc</i> < <i>*hanc:’-</i> ; Andi <i>hinc:’u</i> , <i>Axwax i^hc:’o</i> < <i>*hinč:’-</i>	Avar <i>nuc:’á</i> , Lak <i>nuz</i> < <i>*nvc:’-</i>	NS 218

(continued)

* (c)vcvc-	* (c)vcv/c-	* (c)vcvc/c-	Nikolayev & Starostin 1994
3. * <i>hvmvs(-)</i> 'grouse'	Avar <i>fansá</i> , Inxoqwar <i>eʷzo</i> < * <i>fvms-</i>	Chamalal <i>mus:ija</i> , Archi <i>mus:al</i> < * <i>mvs-</i>	NS 225, 530
4. * <i>hvnvc:-</i> 'honey'	Avar <i>hoc:ó</i> , Andi <i>hunc:i</i> < * <i>hvinc:-</i>	Tsez <i>nuci</i> , Hunzib <i>nucu</i> < * <i>nuc:ə</i>	NS 824
5. * <i>hvrvc:-</i> 'house'	Inxoqwar <i>aʷq</i> , Andi <i>haq:u</i> , Botlix <i>hanq:u</i> < * <i>hanq:-</i> < * <i>harq:-</i>	Avar <i>ruq:ʷ</i>	NS 522
6. * <i>hvlʷn-</i> 'bottom'	Andi <i>hilmu</i> , <i>hinlʷu</i> , Bezhta <i>oʷlʷo</i> < * <i>hvnʷl-</i> < * <i>hvlʷn-</i>	Avar <i>tʷinu</i> , Lak <i>čʷan</i> < * <i>lʷn-</i>	NS 590
7. * <i>bvlʷv-</i> Avar <i>betʷér</i> 'head'	Hinuq <i>miʷlʷu</i> , Bezhta <i>möʷlʷo</i> 'beak' < * <i>bvnʷlʷv</i> < * <i>bvrʷlʷv</i> < * <i>bvlʷlʷv-</i> ; Batsbi <i>marʷo</i> 'nose'	Aghul <i>kʷil</i> , Tabassaran <i>kʷul</i> 'head', Archi <i>li-li-ʷlʷ</i> 'under the head' < * <i>lʷlʷ-</i>	NS 1041
8. * <i>hvk:ʷl-</i> 'mouse': Lak <i>uʷkʷulu</i>	Bezhta <i>aʷqʷo</i> , Avar <i>ʷunk:ʷ</i> , Andi <i>hink:ʷu</i> < * <i>hvnk:ʷv-</i> < * <i>hvlk:ʷv-</i> < * <i>hvk:ʷlv-</i>	Lak <i>kʷulu</i> , Aghul, Tabassaran <i>qʷul</i> < * <i>k:ʷl</i>	NS 523, 935

Paradigms that show this alternation are not preserved in attested East Caucasian languages: they are reconstructed on the basis of a comparison of lexical cognates. These complex paradigms were almost completely replaced by more regular paradigms in the individual languages, by the generalization of one of the two stems, as the examples show. Only very rarely do we find fossils such as Lak *barz*, genitive *zurul* 'moon' or Asax Tsez *uži*, ergative *žojä* 'boy' (cf. Imañshvili 1963: 73).

As examples 6, 7 and 8 indicate, if syncope of the second syllable resulted in a sequence of obstruent + resonant, metathesis occurred to the phonotactically admissible sequence resonant + obstruent. Recall that metathesis was also assumed to have affected the word for 'camel' in Nakh, where **akmil-* became Proto-Nakh **amkil-*. This is important because it shows how well the word for 'camel' was integrated in prehistoric East Caucasian morphophonology to all intents and purposes, it behaves like a native word. Suppose that it was exactly that: Proto-Nakh **amkil* < **akmil* may in that case reflect a Proto-East-Caucasian oblique **ikmil-*, which was the oblique stem (with pretonic unrounding of **ɔ* to **ɪ*) to an absolutive **ɔkmil*, just like the Nakh word for 'glass', **ingal-* ~ **angal-* reflects Proto-East-Caucasian oblique **hingVl/r-* to an absolutive **hɔngV(l/r)* (section 5.1.3).

Against this background it may be more than accidental that example 3, **ʕvmvs(:)*-, whose representatives mean ‘grouse, black grouse, *ular* (Caucasian turkey)’ looks like the European substratum word **amsl-*, **mesal-* ‘blackbird’.

7 Conclusions and outlook

We have seen that Caucasian languages possess a word for ‘camel’ that is so similar to Semitic **gam(a)l-* and Greek κάμηλος that it is probably related. A detailed reconstruction of this word in the Nakh group of East Caucasian reveals a Proto-Nakh paradigm **imkil*, oblique **amkil-*, **amkalV*, which goes back to pre-Proto-Nakh **imkil* or **amkil*, in a paradigm with accentual mobility. In order to be able to connect this with Georgian *aklemi* and with the Semitic and Greek forms, the original position of the **m* was probably after rather than before **k*: **ikmil* or **akmil* (section 5.5). The latter relates to Semitic **gam(a)l-* and Greek κάμηλος according to a pattern that is familiar from a series of European substratum words, such as **amsl-* ~ **mesal-* ‘blackbird’ and **arud-* ~ **raud-* ‘ore’ (section 6). While the latter substrate is implicated in the initial spread of agriculturalists from Asia Minor into Europe from the seventh millennium BCE onwards, the word for ‘camel’ is unlikely to be that old: the domesticated camel reached Eastern Europe and the Near East probably no earlier than the third millennium BC.

The morphophonological pattern observed in **amsl-* ~ **mesal-* and **akmil-* ~ **gam(a)l-*, **kame:l-* has a close counterpart in the morphophonology of the native East Caucasian lexicon, which may indicate that the language that donated the European substrate lexemes was of East Caucasian stock (section 6). The word for ‘camel’ is either a native East Caucasian lexeme, which found its way into Greek and Semitic together with the domesticated Bactrian camel (scenario 2 of section 4); or it is of Semitic origin (and in that case probably originally denoted the dromedary) and spread into East Caucasian so early that it was completely nativized according to the rules of East Caucasian morphology (scenario 1 of section 4). The latter scenario is faced with a serious problem, however. In East Caucasian, the initial syllable that now appears and then disappears in examples like Tsez. *ac* ‘door’ < **anc:*’- < **ʕanVc:*’-, Avar. *nuc:*’á ‘door’ < **nVc:*’- (see the examples in the table in section 7) is not a prefix but the initial syllable of the root of the lexemes in question. Hence, if the connection of East Caucasian **ʕVmVs(:)* ‘grouse’ with European substrate **amsl-* ~ **mesal-* ‘blackbird’ is correct, the **a-* in the latter is not a prefix either. Nor is the **a-* in the word for ‘camel’, **akmil-* ~ **gam(a)l-*, **kame:l-*. If that is so, and if ‘camel’ is a loanword from Semitic into East Caucasian (scenario 1 of section 4), how, then, would it have acquired the initial **a-*,

which is absent from Semitic? The converse is more likely: an East Caucasian alternating **akmil-* ~ **kVmVl-* was borrowed into Semitic and Greek on the basis of the latter root form. Hence it is necessary to come out in favour of scenario 2.

If ‘camel’ and/or the European substrate lexicon referred to above are indeed of East Caucasian origin, this has an important consequence. The East Caucasian family is attested in the northeastern Caucasus and is not known to have ever resided elsewhere. However, if the language of the European substrate lexicon is East Caucasian *and* the language of the first European farmers, that must mean that East Caucasian was more widely distributed in the Near East by the early Neolithic. Similarly, if ‘camel’ is a native East Caucasian lexeme, the fact that it became widespread in Semitic and ended up in Greek implies that East Caucasian could be found somewhere outside the Caucasus area, and more precisely even as far east as the area between Turkmenistan and eastern Turkey, by the third millennium BCE. And the fact that East Caucasian is nowadays spoken only in the Caucasus would put it on a par with Ossetic and Armenian, as a representative of a once widespread language family that managed to hold on to life only in the Caucasus while it was replaced by other languages families in the lowlands to the north and south. These are far-reaching conclusions, which require a stronger foundation than can be provided on the basis of our present state of knowledge. Whether the ideas expressed here will ever surpass the status of ideas will depend on progress on the reconstruction of East Caucasian, but a tantalizing piece of material may support the idea, however tentatively: Tocharian, an Indo-European language that was spoken in oases along the Silk Road in present-day Xinjiang, possesses a word for ‘elephant’ that is formally similar to Nakh **akmil*, viz. Tocharian B **onkolmo*, Tocharian A *onkalām* < Proto-Tocharian **enkelmo*.¹⁶ For the plausibility of a semantic transition from ‘camel’ to ‘elephant’, compare the Sumerian neologism for ‘Bactrian camel’, **am.si.ħar.an**, which literally means ‘elephant of the road/caravan’ (see section 3); also note the plausible etymological connection between Gothic *ulbandus* ‘camel’ and Greek ἑλέφας, -αντ- ‘elephant’ (e.g. Lehmann 1986: 375).

Finally, to return to the beginning of this study, pre-Proto-Nakh **akmil-* and its presumed co-existence with a form without **a-*, **kamil-*, may be relevant to explaining why Greek κάμηλος has an initial *k-* and a front vowel in the second syllable, issues that are both left unexplained by the idea that the Greek word was borrowed from Semitic **gam(a)l-*.

¹⁶ I am indebted to Guus Kroonen for pointing me to this form. See Adams 1999: 113 for attempts to provide the word with an Indo-European etymology. The Tocharian word for ‘camel’ is hiding behind the Tocharian B adjective *partāktaiñe* ‘pertaining to a camel’, which lacks a convincing etymology (Adams 1999: 358–59).

References

- Adams, Douglas Q. 1999. *A Dictionary of Tocharian B*. Amsterdam/Atlanta: Rodopi.
- Aliroev, Ibragim Junusovich. 1975. *Sravnitel'no-sopostavitel'nyj slovar'otraslevoj leksiki čečenskogo i ingušškogo jazykov i dialektov* [*Comparative Dictionary of the Word-Field Lexicon of the Chechen and Ingush Languages and Dialects*]. Maxačkala: Čečeno-Ingušskoe Knižnoe Izdatel'stvo.
- Almathen, Faisal et al. 2016. Ancient and modern DNA reveal dynamics of domestication and cross-continental dispersal of the dromedary. *Proceedings of the National Academy of Sciences*, June 14, 2016. 6707–6712.
- Bauer, Anna. 2020. Hieroglyphic Luwian /*kamara?/, ka-mara/i-* '(a commodity)'. In *EDiAنا (Digital Philological-Etymological Dictionary of the Minor Ancient Anatolian Corpus Languages)*, Beta-version 0.2, consulted 1-1-2022 Available at <https://www.ediana.gwi.uni-muenchen.de/dictionary.php?lemma=988>.
- Chantraine, Pierre. 1968. *Dictionnaire étymologique de la langue grecque*. Paris: Klincksieck.
- Desheriev, Junus D. 1960. *Sovremennyj čečenskij literaturnyj jazyk, časť I, Fonetika* [*The Modern Chechen Literary Language, Part I, Phonetics*]. Groznyj: Čečeno-Ingušskoe Knižnoe Izdatel'stvo.
- Gagua, R.R. 1961. Sklonenie suščestvitel'nyx v Batsbijskom jazyke [The inflection of substantive nouns in the Batsbi language]. In E.A. Bokarev (ed.), *Voprosy izučeniya iberijsko-kavkazskix jazykov* [*Questions of the Study of the Ibero-Caucasian Languages*], 72–88. Moscow: Izdatel'stvo Akademii Nauk.
- Giginejshvili, Bakar K. 1977. *Sravnitel'naja fonetika dagestanskix jazykov* [*Comparative Phonetics of the Dagestan Languages*]. Tbilisi: Izdatel'stvo Tbilisskogo Universiteta.
- Heide, Martin 2011. The domestication of the camel: Biological, archaeological and inscriptional evidence from Mesopotamia, Egypt, Israel and Arabia, and literary evidence from the Hebrew bible. *Ugarit-Forschungen* 42. 331–384.
- Holisky, Dee Ann & Rusudan Gagua 1994. Tsova-Tush (Batsbi). In Rieks Smeets (ed.), *The Indigenous Languages of the Caucasus*, Vol. 4: *The North East Caucasian Languages*, Part 2, 147–212. Delmar/ New York: Caravan Books.
- Imnajshvili, David S. 1963. *Didojskij jazyk v sravnenii s ginuxskim i xvaršijskim jazykami* [*The Dido Language in Comparison with the Ginukh and Khwarshi Languages*]. Tbilisi: Izdatel'stvo Akademii Nauk Gruzinskoj SSR.
- Imnajshvili, David S. 1977. *Istoriko-sravnitel'nyj analiz fonetiki naxskix jazykov* [*Historical-Comparative Analysis of the Phonetics of the Nakh Languages*]. Tbilisi: Mecniereba.
- Iversen, Rune & Guus Kroonen. 2017. Talking Neolithic: linguistic and archaeological perspectives on how Indo-European was implemented in Southern Scandinavia. *American Journal of Archaeology* 121(4). 511–525.
- Jakovlev, Nikolaj F. 1960. *Morfologija čečenskogo jazyka* [*Morphology of the Chechen Language*]. Groznyj: Čečeno-Ingušskoe Knižnoe Izdatel'stvo.
- Kadagidze, Niko & David Kadagidze. 1984. *C'ova-tušur-kartul-rusuli lek'sikoni* [*Tsova-Tush-Georgian-Russian Lexicon*]. Tbilisi: Mecniereba.
- Khalilova, Zaira. 2009. *A Grammar of Khwarshi*. Utrecht: LOT.
- Kibrik, Aleksandr E. & Sandro V. Kodzasov 1990. *Sopostavitel'noe izučenie dagestanskix jazykov* [*Comparative Study of the Dagestan Languages*]. Moscow: Izdatel'stvo Moskovskogo Universiteta.
- Kirtcho, Lyubov B. 2009. The earliest wheeled transport in southwestern Central Asia: new finds from Altyn Depe. *Archaeology, Ethnology & Anthropology of Eurasia* 37(1). 25–33.

- Klimov, Georgij A. & Madzhid Sharipovich Xalilov. 2003. *Slovar' kavkazskix jazykov. Sopostavlenie osnovnoj leksiki [Dictionary of the Caucasian Languages. Comparison of the Basic Lexicon]*. Moscow: Vostočnaja Literatura.
- Kohl, Philip L. 2007. *The Making of Bronze Age Eurasia*. Cambridge: Cambridge University Press.
- Komri, Bernard & Madzhid Sharipovich Xalilov, Zaira Xalilova. 2015. *Grammatika bežtinskogo jazyka [Grammar of the Bezhta Language]*. Leipzig: Max Planck Institut for Evolutionary Anthropology, and Maxačkala: Institut Jazyka, Literaturny i Iskusstva im. G. Cadasy, Dagestanskij Naučnyj Centr Rossijskoj Akademii Nauk.
- Kossmann, Maarten. 2005. *Berber Loanwords in Hausa*. Cologne: Rüdiger Köppe.
- Kretschmer, Paul. 1892. Zum ionisch-attischen Wandel von $\tilde{\alpha}$ in η . *Zeitschrift für vergleichende Sprachforschung* 31. 285–296.
- Kroonen, Guus. 2012. On the etymology of Greek $\alpha\gamma\lambda\iota\varsigma$ and $\gamma\acute{\epsilon}\lambda\iota\varsigma$ 'garlic'. *Journal of Indo-European Studies* 40. 289–299.
- Lehmann, Winfred P. 1986. *A Gothic Etymological Dictionary*. Leiden: Brill.
- Mallory, James P. & Douglas Q. Adams (eds.). 1997. *Encyclopedia of Indo-European Culture*. London/Chicago: Fitzroy Dearborn Publishers.
- Masson, Émilie. 1967. *Recherches sur les plus anciens emprunts sémitiques en grec*. Paris: Klincksieck.
- Matsiev, Axmet G. 1961. *Cečensko-russkij slovar' [Chechen-Russian Dictionary]*. Moscow: Gosudarstvennoe izdatel'stvo inostrannyx i nacional'nyx slovarej.
- Militarev, Alexander & Leonid Kogan. 2005. *Semitic Etymological Dictionary. Volume II Animal Names*. Münster: Ugarit-Verlag.
- Nichols, Johanna. 1994. Chechen. In Rieks Smeets (ed.), *The Indigenous Languages of the Caucasus, Vol. 4: The North East Caucasian Languages, Part 2, 1–77*. Delmar/New York: Caravan Books.
- Nichols, Johanna. 1997. Chechen phonology. In A. S. Kaye (ed.), *Phonologies of Asia and Africa (including the Caucasus)*, 941–971. Winona Lake, IN: Eisenbrauns.
- Nichols, Johanna. 2003. The Nakh-Daghestanian consonant correspondences. In Dee Ann Holisky & Kevin Tuite (eds.), *Current Trends in Caucasian, East European and Inner Asian Linguistics. Papers in Honor of Howard I. Aronson*, 207–264. Amsterdam/Philadelphia: John Benjamins.
- Nichols, Johanna. 2004. *Ingush-English and English-Ingush Dictionary*. London/New York: Routledge.
- Nichols, Johanna. 2011. *Ingush Grammar*. Berkeley/Los Angeles/London: University of California Press.
- Nichols, Johanna & Arbi Vagapov. 2004. *Chechen-English and English-Chechen Dictionary*. London/New York: Routledge.
- Nikolayev, Sergei L. & Sergei A. Starostin. 1994. *A North Caucasian Etymological Dictionary*. Moscow: Asterisk Publishers.
- Potts, Daniel T. 2004. Camel hybridization and the role of *Camelus bactrianus* in the ancient Near East. *Journal of the Economic and Social History of the Orient* 47. 143–165.
- Puhvel, Jaan. 1984. *Hittite Etymological Dictionary*. Vol. 1: *Words beginning with A*, Vol. 2: *Words beginning with E and I*. Berlin/New York/Amsterdam: Mouton Publishers.
- Salimov, Xangerej S. 2010. *Gagatliński govór andijskogo jazyka [The Gagatl Dialect of the Andi Language]*. Maxačkala: IJaLI.
- Schrijver, Peter. 2018a. The origin of vowel alternation in Avar-Andi-Dido (North-East Caucasian), with special reference to the Dido languages. *International Journal of Diachronic Linguistics* 15. 199–223.
- Schrijver, Peter. 2018b. Talking Neolithic: the case for Hatto-Minoan and its relationship to Sumerian. In Guus Kroonen, James P. Mallory & Bernard Comrie (eds.), *Talking Neolithic (Proceedings of the workshop on Indo-European origins held at the Max Planck Institute for Evolutionary*

- Anthropology, Leipzig, December 2–3, 2013), 336–374. Washington: Institute for the Study of Man.
- Schrijver, Peter. 2021. A history of the vowel systems of the Nakh languages (East Caucasian), with special reference to umlaut in Chechen and Ingush. *Languages of the Caucasus* 5. 84–151.
- Tischler, Johannes. 1977. *Hethitisches etymologisches Glossar*, Lieferung 1. Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft.
- Uerpmann, Hans-Peter & Margarethe Uerpmann. 2002. The appearance of the domestic camel in southeast Arabia. *Journal of Oman Studies* 12. 235–260.
- Xalilov, Madzhid Sharipovich 1995. Bezhtinsko-russkij slovar' [Bezhta-Russian Dictionary]. Maxachkala: Institut Jazyka, Literatury i Iskusstva im. G. Cadasy.
- Xalilov, Madzhid Sharipovich 1999. *Cezsko-russkij slovar'* [Tsez-Russian Dictionary]. Maxachkala: Institut Jazyka, Literatury i Iskusstva im. G. Cadasy.
- Xalilov, Madzhid Sharipovich & Isak Abdulvaxidovich Isakov. 2005. *Ginuxsko-russkij slovar'* [Ginukh-Russian Dictionary]. Makhachkala: Institut Jazyka, Literatury i Iskusstva im. G. Cadasy.
- Yakubovich, Ilya. 2016. Some transitive motion verbs and related lexemes in Late Luwian. *Indogermanische Forschungen* 121. 69–92.

