Johanna Nichols

University of California, Berkeley

# The history of an attractor state: Adventitious *m* in Nakh-Daghestanian pronominals

## 1.  Introduction

Juha Janhunen's reliable reconstruction of protolanguages (Janhunen 1981, 1982), detection of plausible traces of long-range connections between proto-languages (e.g. Janhunen 1996), and arguments against other proposed deep connections (e.g. Janhunen 1996, 2001) carry several imperatives for a linguist working on pronominal systems in Eurasia. Pronouns have often figured decisively in demonstrations of genealogical relatedness (e.g. Sapir 1913 and Haas 1958 on Algic, Greenberg 1960 and Newman 1980 on Afroasiatic). They have also figured prominently in long-range claims that have not been demonstrated (Nostratic, Eurasiatic, Amerind; not demonstrated in the sense that the best evidence offered by proponents has been shown to be insufficient). They present a thorny problem for genealogical heuristics because pronouns are known to be stable vocabulary in general, yet they often undergo idiosyncratic sound changes and morphological changes that readjust paradigms considerably (Meillet 1925); and unless plural forms are independent of singular forms there may not be enough distinct roots in a pronoun system for statistical demonstration of relatedness (Nichols & Peterson 1996). Despite Janhunen's careful demonstration that many of the putative cognates among Turkic and Mongolic are not in fact valid cognates (Janhunen 1996, 2001, 2003a-b, in press), the personal pronouns in these and other supposed Nostratic language families remain a striking similarity requiring an explanation.

The present paper attempts a step toward such an explanation, continuing my own recent work on Eurasian pronominals (Nichols in press; also Nichols & Peterson 2011, 2005, 1996, Nichols 2001; the typological and methodological background are laid out in Nichols & Peterson 2011, 2005) which argues that the resemblant pronominals with first person *m* and second person coronal obstruents, common to Uralic, Turkic, Tungusic, Mongolic, Indo-European, Yukagir, and Kartvelian, far from being an ancient inheritance,[1] are what is known as an

---

1.    In any event first person *m* is known to be secondary in Turkic, Tungusic, and Mongolic (e.g. Janhunen 2003a:18, Johanson 1998: 112).

*attractor state* in complexity theory: a form or alloform which arises more easily than it is lost and therefore is more prone to expand over time than to retract. (More precisely, the probability of expansion is greater than the possibility of loss. The probability differential need not be great; even if the attractor state has only a small edge over its competitors, it will inexorably expand in the long run. In pronominals attractor states include nasal consonants and especially [m], probably because it is both basic and easily distinguished, and to a lesser extent other basic and high-frequency consonants. Paradigmatic sets are moderately likely to innovate such forms and fairly unlikely to lose them, so they tend to expand over time.) The expansion of first person *m* in eastern Eurasia was furthered by the distinctive sociolinguistics of the late prehistoric and protohistoric eastern steppe, which involved close and changing contact with back-and-forth language shift, all of which presents speakers with more possibilities to hear attractor states and increases the opportunities for an attractor state to expand.

Here I trace the rise and spread of *m* and systems with *m* and a coronal obstruent in the Nakh-Daghestanian (ND) family of the eastern Caucasus, a family which, though far from the eastern Eurasian steppe, has some general sociolinguistic resemblances to the steppe situation accompanied by very general typological resemblances to the pronominal systems of eastern Eurasia.[2] Little has been done on Nakh-Daghestanian pronouns. Schrijver 2009 reconstructs personal pronouns for the Avar-Andic-Tsezic subbranch and looks to the Lezgian subbranch to suggest a Proto-ND reconstruction, but his ongoing work has not yet covered the Lak and Dargwa subbranches or the Nakh branch. Schulze 1999 considers chiefly ergative and nominative (which are syncretic in many Daghestanian languages) and shows that ergative morphology reconstructs for most of the daughter branches but the morphology is not cognate between the branches and cannot be reconstructed for Proto-ND. Gigineishvili 1977 identifies consonant correspondences among some of the Daghestanian pronominals but does not attempt a full reconstruction for Daghestanian and does not consider Nakh at all. Nikolayev and Starostin 1994 assume Nakh-Daghestanian and West Caucasian (a different indigenous Caucasian family) are related and reconstruct a pronominal system that unites them. Schrijver and Schulze do modern analysis involving both phonology and morphology; Gigineishvili seeks only phonological correspondences and does not reconstruct whole forms; Nikolayev and Starostin reconstruct forms and generally seek exclusively phonological solutions to correspondences. But the pronominal systems of the ND branches display frequent and to my mind striking evidence for morphological readjustments and processes of analogy, some of them also shared by interrogative pronouns. Therefore this paper attempts a primarily morphological analysis of the singular personal pronouns, the interrogative pronouns 'who' and 'what', and

---

2.     Nakh-Daghestanian, also known as East Caucasian, is a very old family with about 40 daughter languages (including mutually unintelligible dialects) in deeply divided branches. It bifurcates initially into Nakh and Daghestanian, and Daghestanian further divides into five to seven branches. Tables 1–3 show the branch names and language names for the languages used here.

The history of an attractor state:
Adventitious *m* in Nakh-Daghestanian pronominals
263

the main interrogative adverbs ('when', 'where', 'how', etc.). In a nutshell, the occasional consonant with an originally more restricted distribution has expanded across one or another pronominal subset or paradigm in one or another daughter branch, and *m* is prone to such expansions.

## 2.   Personal pronouns: 1sg, 2sg.

Relevant forms are shown in Table 1.[3] Of main interest is the first consonant in each form (bold in the tables). The discussion is limited to first and second person forms because ND languages generally lack dedicated third person pronouns; various demonstratives are used instead. Only singular forms are surveyed because in ND languages the plural pronouns are so often based on singulars in one way or another that they contribute little independent evidence. The Nakh exclusive is cognate to the Daghestanian inclusive, and the inclusive/exclusive distinction is lost in some branches, complicating any discussion of the first person plural pronouns.

The discussion to follow deals almost entirely with consonants and with their morphological rather than phonological nature: occasionally the phonetic value of a proto-phoneme requires discussion, but for the most part the only point at issue is whether this or that root consonant is shared by other forms in the paradigm and/or by other paradigms. Also occasionally at issue is whether the root morpheme starts out with *CV-* or *VC-*. Alternations between CV- and VC- shapes occur in noun and pronoun paradigms of several of the daughter languages, with CV- usually in the nominative[4] and VC- in one or more oblique forms. In the first person nominative, Ingush *so* 'I, me', Avar *dun*, Tsex *di*, Lak *na*, Kubachi Dargwa *du*, Lezgi *zun*, and most others have CV shape. The Nakh ergatives *aaz, as* have VC shape, as do several Lezgian genitives: Tabassaran *jas, jiz*, Aghul *jaz*, Rutul *iz*, Tsakhur *jiz-*, Archi *C=is*, Udi *bez*. In these genitives, the Archi C= is a gender prefix agreeing with the head (possessed) noun,

---

3.      The forms come straight from my database, which uses exclusively the Latin letters used in English orthography, without special symbols or diacritics. Digraphs are used for some phones: *ch, sh*, etc. are as in English; *tl* is a lateral affricate. *c* is [ts]; *y* is [ɨ]; *ë* is schwa. *x* and *gh* are uvular fricatives (there are no true velar fricatives in the data here). 9 = pharyngeal, 7 = glottal stop (but ejectives are marked with an apostrophe: c', q', etc.). Geminates and long vowels are written with doubled letters (geminate digraphs repeat only the first letter, e.g. *cch* rather than *chch*). Capital N = nasalization; G = uvular stop. Following the convention used by Kibrik and Kodzasov 1990, the equals boundary segments off gender agreement markers. In the spellings =*C* and *C*=, C indicates the varying gender marker (which is usually a consonant). These conventions guarantee searchability from any platform and stability across software changes and with non-Unicode-ready software. The very simple syllable canon of Proto-Nakh-Daghestanian and most of its daughters makes digraph spellings of consonants possible, since there are very few consonant clusters in native vocabulary.

4.      I use *nominative* rather than *absolutive* for the (usually unsuffixed) citation form, since by no means all verbs in the daughter languages have ergative (S=O≠A) alignment. In fact in most daughter languages, and likely in the protolanguage, simple verbs were a closed class and new verbs were formed with phrasal predicates using light verbs. Since most of the light verbs are transitive they assign the ergative, even if the phrasal predicate has only one argument (e.g. Ingush *aaz nab ju* [I-ERG sleep do-PRES] 'I sleep'). See also Creissels 2009 for more reasons to use the term *nominative* even for ergative languages.

| Branch, Language | Dialect | 1 sg Nom. | Erg. | Gen. | Dat. | 2 sg Nom. | Erg. | Gen. | Dat. |
|---|---|---|---|---|---|---|---|---|---|
| **Nakh:** | | | | | | | | | |
| Ingush | | **so** | aa**z** | **s**y | **s**uo- | **h9**o | a**h9** | **h9**a | **h9**uon- |
| Chechen | lowland | **so** | a**s** | **s**aN | **s**uo- | **h9**o | a**h9** | **h9**aN | **h9**uon- |
| Batsbi | | **so** | a**s** | **s**eN | **s**o- | **h9**o | a**h9** | **h9**eN | **h9**o- |
| **Avar-Andic:** | | | | | | | | | |
| Avar | standard | **d**un | **d**i- | **d**i- | **d**i- | **m**un | du- | du- | du- |
| Avar | Antsux | **d**un | **d**i- | **d**in | **d**i- | **m**un | du- | du- | du=C=e |
| Andi | | **d**in/**d**en | **d**in | **d**i- | **d**i=C | **m**in/**m**en | **m**in/**m**en | du= | du- |
| Akhvakh | northern | **d**ene | **d**e- | **d**i=C | **d**i- | **m**ene | **m**e- | du | du- |
| Chamalal | L. Gakv. | **d**iiN | **d**e: | **d**i= | **d**i- | **m**i: | **m**in | du=C | du- |
| Chamalal | Gigatli | **d**e=C | **d**ennu | **d**i=C | **d**i- | **m**i=C | **m**inni | du=C | du- |
| Tindi | Lower | **d**e | **d**i | **d**i=C | **d**i- | **m**e | **m**i | du- | du- |
| Botlikh | | **d**en(i) | ishkur | **d**i=C | **d**i- | **m**in(i) | **m**in(i) | du=C | du- |
| Godoberi | | **d**en | **d**en | **d**i=C | **d**i- | **m**in | **m**in | du=C | du- |
| Karata | | **d**en | **d**en-a | **d**i=C | **d**i- | **m**en | **m**en-a | du=C | du- |
| Karata | Tukita | **d**ini | **d**in-di | **d**i=C | **d**i- | **m**ini | **m**in-di | di=C | di- |
| Bagwalal | | **d**e(N) | **d**en | **d**i=C | **d**i- | **m**eN | **m**en | du=C | du- |
| **Tsezic:** | | | | | | | | | |
| Tsez | | **d**i | **d**i- | **d**ej | **d**är- | **m**i | **m**i | deb- | deb- |
| Xvarshi | Inxokvari | **d**o | **d**e | **d**i- | **d**i- | **m**o | **m**e | dub- | dub- |
| Hinuq | | **d**e | **d**e | **d**i | **d**i- | **m**e | **m**e | debe | debe- |
| Hunzib | | **d**ë | **d**ë | **d**ije | **d**i7i | **m**ë | **m**ë | dëbë | dibi |
| Bezhta | Tljadal | **d**o | **d**o | **d**i- | **d**i- | **m**i | **m**i | di-**b**o | du- |
| **Lak** | standard | na | **ttu**- | **ttu**- | **ttu**- | ina | **wi**- | **wi**- | **wi**- |
| **Dargwa:** | | | | | | | | | |
| Dargi | standard | **n**u (du) | **n**uni | **d**i- | **n**ab | **h9**u | **h9**uni | **h9**e-la | **h9**e-d |
| Chiragh | Chiragh | **d**u | **d**i- | **d**i- | **d**amii | **9**u | **9**i- | **9**i- | **9**at |
| Kubachi | Kubachi | **d**u | **d**u-di- | **d**i- | **d**ammi-/ **d**am | u | u-di- | i- | itti- |
| Mehweb | Mehweb | **n**u | **n**u- | **d**i- | **n**a-b | **h9**u | **h9**u- | **h9**a- | **h9**a-d |
| Sanzhi-Ic'ari | Ic'ari | **d**u | **d**ul | **d**ila | **d**am | u | ul | ila | att |
| Xaidaq' | Xaidaq' | **d**u | **d**u-li | **d**i-la | izh; **d**am | i | i-li | i-la | e-t |
| **Lezgian:** | | | | | | | | | |
| Lezgi | standard | **z**un | **z**a | **z**i | **z**a-z | **w**un | **w**una, na | **w**i | **w**a-z |
| Lezgi | Axty | **z**yn | **z**a | **z**y | **z**a-z | **w**yn | **w**yn- | **w**i | **w**a- |
| Tabassaran | North | izu | izu | ja**s** | izu- | iwu | iwu | jaw | iwu- |
| Tabassaran | South | uzu | uzu | ji**z** | uzu- | uwu | uwu | jaw | uwu- |
| Aghul | Burschag | **z**un | **z**un | ja**z** | **z**as | **w**un | **w**un | jawir | **w**as |
| Aghul | Richa | **z**un | **z**ash | **z**e | **z**us | **w**un | **w**ash | **w**e | **w**as |
| Aghul | Burkixan | **z**un | **z**un | **z**e | **z**is | **w**un | **w**un | **w**e | **w**as |
| Aghul | Fite | **z**un | **z**u**n** | **z**it | **z**as | **w**un | **w**un | **w**it | **w**as |
| Rutul | | **z**y | **z**a-d | iz-dy | **z**a-s | **w**y | **w**a- | **w**y- | **w**a- |
| Tsaxur | Mixik | **z**y | **z**a-sse | ji**z**-y=n | **z**a-s | **gh**u | **w**a- | **w**as-y= | **w**a- |
| Kryz | | **z**yn | **z**yn | **z**ä | **z**äs | **v**un | **v**un | väs | **9**a= |
| Budukh | | **z**yn | **z**yn | **z**ëz | **z**o | **v**yn | **v**yn | **v**o | **v**ëz |
| Archi | | **z**on | **z**a- | C=i**s** | C=e**z** | un | un | **w**it | **w**a- |
| Udi | | **z**u | **z**u | be**z** | **z**a | hun | hun | **v**i | **v**a |
| Xinalug | | **z**y | jä | i / e | a**s** | **v**y | **v**a | **v**- | o**X**= |

Table 1. Personal pronouns[5]

and the Udi *b-* and Tabassaran and Aghul *j-* are frozen gender markers.[6] Some daughters have leveled out such alternations by generalizing one or the other shape. In the first person pronouns, many languages generalize the CV- shape to all forms (e.g. Avar, with nominative *dun* and general oblique stem *di-*), but a few generalize the VC- shape (e.g. Tabassaran [Lezgian], with nominative *izu*, general oblique *izu-*, genitive *jas*).

1sg: All Daghestanian branches regularly reflect PND *d (except that Archi and Udi irregularly generalize a non-initial reflex: Gigineishvili 1977: 76). Nakh *s* also appears to reflect *d. I have no clear examples of PND *d in Nakh other than the gender prefix *d-*, which has been kept regular, transparent, and free of allomorphy by paradigmatic pressure. However, strong evidence that Nakh *s* reflects PND *d comes the fact that Daghestanian has *d- in both the first person singular pronoun and a major ergative allomorph (Gigineishvili 1977: 76), and Nakh has *s* in both of these forms. (*d, like any reconstruction, labels a set of correspondences and should not be read as a phonetic transcription. It is entirely possible that its actual phonetic value was a stop or affricate in at least some contexts.)[7]

CV(-) forms predominate; VC forms are found in the Nakh ergative and in Lezgian (obliques in Burschag Aghul, Rutul, Tsakhur, Archi, Udi, and Xinalug; the whole paradigm in Tabassaran).[8]

The initial *n- of the Lak nominative *na* and the Mehweb and standard Dargi obliques *nab* seem to be due to distant assimilation of the regular initial *d-*. Or perhaps this is nasalization metathesis; the dative forms in *dam-* in the other Dargwa languages are probably conservative and standard Dargi and Mehweb *nab* innovative. Distant nasal assimilation is fairly common in the Daghestanian languages (for Avar-Andic see Gudava 1959).

Thus a reconstruction of a single root consonant *d is straightforward, though not always regular. Though first person singular pronouns are quite prone, cross-linguistically, to suppletive stem allomorphy, there seems to have been none in Proto-ND or the early branches. Only when distant nasal assimilation becomes unproductive (as in Lak *na, ttu-* and Dargi *nu, di-*) do stems in the daughter languages begin to look suppletive.

---

5.     Sources: Kibrik & Kodzasov 1990 for most languages; also Nichols 2011 (Ingush), Nichols & Vagapov 2004 (Chechen), Holisky 1994 (Batsbi), Kibrik ed. 2001 (Bagwalal), Kibrik et al. 1996 (Godoberi), Magometov 1963 (Kubachi), 1982 (Mehweb), Sumbatova & Mutalov 2003 (Ic'ari), Temirbulatova 2004 (Xaidaq), Haspelmath 1993 (standard Lezgi).

6.     Lezgi, Aghul, and Udi have lost gender entirely but retain some fossilized gender prefixes. In languages that preserve gender it is not uncommon for one or another word to lose agreement and reanalyze a gender prefix as a root-initial consonant. In most and probably all of the languages, some words in the agreeing classes (chiefly verbs and adjectives) agree and some do not, so it is a simple matter for a verb to stop agreeing and freeze one or another gender marker as its root-initial consonant.

7.     Schulze 1999 reconstructs *z and regards the Avar-Andic-Tsezic forms with *d-* as not cognate to the rest. Gigineishvili 1977: 76–7 reconstructs *d and sees the Lezgian reflex *z* as the regular outcome of a Proto-Lezgian or early Lezgian lenition in initial position before an unstressed syllable.

8.     The suppletive Botlikh ergative is isolated in the family and in its paradigm, and its origin is unknown (Schulze 1999: 104). In Lezgian dative forms like Lezgi *za-z* the *-z* is a productive dative ending and not a reduplicated root consonant. The similar-looking final consonants of Archi dative *C=ez* and Xinalug *as* are not dative endings but must be the root consonants.

**2sg:** Though even less regular, reconstruction of a root consonant seems fairly straightforward here as well. The initial consonants are quite varied: pharyngeal in Nakh, labial in Avar-Andic-Tsezic, no consonant (vocalic initial, usually realized with a prevocalic glottal stop) in Lak and some Dargwa, pharyngeal in other Dargwa, uvular in Tsakhur, /v/ or /w/ in most of Lezgian (the Lezgian reflexes are not entirely regular, but Alekseev 1985: 72 suggests reconstructing Proto-Lezgian *gh$^w$). The stem vowel is often rounded. I would reconstruct a uvular, probably the /gh/ reflected in Lezgian, since uvulars are prone to turn into pharyngeals (Carlson & Esling 2003 for Wakashan and Salish; as an ND example, in one of the words for 'grain' Avar *bu9a* corresponds to Akhvakh *muq'a*, Bezhta *müq'e,* and Archi *buq'*) but the reverse is rare if it occurs at all. Similar correspondences occur between Nakh and Chirag Dargwa in 'apple' (Nichols 2003: 263 #57–58), possibly 'eye' (Nichols 2003: 264 #62); Nakh and Xinalug in 'grind' (Nakh *\*ah9-*, Xin. *cXu-v*) and its derivative 'mill' (Nakh *\*h9air-*, Xin. *zoX*).

This initial consonant was evidently labialized, perhaps only secondarily in assimilation to the round vowel (Nakh *o, Daghestanian *u) that occurs in many of the case forms. This labialized consonant then underwent distant nasal assimilation in Avar-Andic and Tsezic. Whatever the identity of the initial consonant, it was evidently more prone to nasal assimilation than *d or its reflex was, as nasal reflexes are more numerous in the second person than in the first. In the first person they occur only in Lak *na* and Dargi nom. *nu*, erg. *nuni*, while in the second person they occur throughout Avar-Andic-Tsezic.

Again, both CV- and VC- forms are attested: CV(-) forms predominate in the nominative, VC- forms in the oblique forms (Nakh ergative, Tabassaran genitive, Xinalug dative).

To summarize, both the first and second person singular pronouns had a single Proto-ND root consonant, 1sg *d and 2sg *gh$^{(w)}$, that occurred in two shapes, CV(-) at least in the nominative and VC(-) in at least one oblique case. They were the only root consonants in their respective paradigms.

In Avar-Andic and Tsezic, the second person nominative *mun* of Avar forms a suppletive paradigm with *du-* in the oblique cases, as do the syncretic nominative-ergative cases of the Andic and Tsezic languages (e.g. Tsez nom.-erg. *mi*, oblique *deb-*). There is no obvious source for the oblique *d-*, other than analogical extension from the first person singular paradigm (where it is the reflex of Proto-ND *d-*). First and second person pronouns are differentiated by different vowels and, in Tsezic, a stem-final *-b* (which is also the initial consonant of the second person plural pronoun, not shown here). Apparently first person *d-* was reanalyzed as a marker of singular pronouns and extended to the second person. Though this is speculative, I imagine that the reason for the extension of first person morphology to the second person was that, after the initial *gh$^w$- had been nasalized in the nominative, the oblique stem with its initial non-nasal (*gh$^w$u-? wu-?*), perhaps in a *VC-* shape (*ugh-? uw-? ub-?*) had come to differ so much from the nominative form that the paradigm was perceived synchronically

as suppletive. Perhaps the /u/ vocalism of the Avar-Andic second person oblique stems is the surviving trace of the original root consonant. It is possible that the Tsezic -*b*- of the oblique second person forms is also a surviving trace of the original *gh$^{(w)}$ in postvocalic position. Whatever the source of the -*b*-, I suggest, still speculating, that first person *d*- was recruited to give enough skeletal weight and/or paradigmatic coherence to the oblique forms of the second person. The modern synchronic result is a suppletive root for the second person singular paradigm and an initial consonant shared by first and second persons.

My reconstruction follows Gigineishvili for 1sg; for 2sg he comments only on the final *-n found in some forms. I believe it is consistent with Schrijver's reconstruction (2009), though not identical as I draw more on Nakh, Lak, and Dargi. It is superficially different from Schulze's (1999), in that he treats Avar-Andic-Tsezic 1sg *d* as non-cognate to the rest and I regard it as irregular or context-specific but cognate. Nikolayev and Starostin 1994 reconstruct for Proto-ND 1sg *zo:, 2sg *wo:, taking the Nakh 1pl inclusive *vai* rather than the 2sg *h9o* to descend from their Proto-ND 2sg. These reconstructions enable them to draw parallels to West Caucasian and Hurrian-Urartian, but I find their posited changes of *z to *d (in Avar-Andic, Tsezic, and Dargwa) and *w to uvular *gh* (Tsakhur), uvular *x* (Xinalug), and pharyngeal *h9* and *9* (Dargwa) non-parsimonious (among other things the family has a number of instances of lenition and little or no fortition) and the change from *2sg to 1pl inclusive highly non-parsimonious.[9]

## 3.   Interrogative pronouns: 'who', 'what'

Table 2 shows the forms for 'who' and 'what' in the nominative, the oblique stem, and a second oblique stem if different. The nominative and oblique stems have different grammatical behavior. It is common in Daghestanian languages to have a single stem for 'who' and 'what', distinguished only by gender suffixes, e.g. Avar *sshi=w* 'who (masc.)', *sshi=j* 'who (fem.)', *sshi=b* 'what' (neuter). (The gender agrees with the antecedent or referent of the pronoun.) This pertains only to the nominative; oblique stems distinguish 'who' from 'what' and for the most part do not mark gender. (Recall that genitives in some languages agree with the possessed head noun; but they do not agree with the referent of the pronoun as the Avar examples just cited do.) The oblique stems are consistent from branch to branch and can be reconstructed for Proto-ND, while nominative stems are varied: they are fairly consistent within branches but show little consistency between branches. That is, oblique stems are considerably more stable

---

9.     A number of the Nikolayev-Starostin ND cognate sets seem to me to contain words that would not be judged cognate on a family-internal reconstruction but are put together because they support a ND-West Caucasian connection. For this reason the ND-internal reconstructions given here do not draw very much on their work.

Oblique stem 2 entered only if different from oblique 1.

| Branch, Language | Dialect | who Nom. | Obl. | Obl. 2 | what Nom. | Obl. | Obl. 2 |
|---|---|---|---|---|---|---|---|
| **Nakh:** | | | | | | | |
| Ingush | | mala | **h9**an | | fy | **s**ien- | |
| Chechen | lowland | mila | **h9**an | **h9**aan- | huN | **st**ie | **st**ien- |
| Batsbi | | meN | **h9**an | | vux | **st**’en- | |
| **Avar-Andic:** | | | | | | | |
| Avar | standard | sshi=C | **lh**i- | | sshi=C | **s**un- | |
| Avar | Antsux | su | **lh**i- | | sib | **s**undu | lhi- |
| Andi | | emi- | **llh**e- | | ebi- | llhun- | |
| Akhvakh | northern | chu- | **llh**o- | | chu- | **ss**uN- | |
| Chamalal | L. Gakv. | im- | **llh**e- | | ed | llhun- | |
| Chamalal | Gigatli | ik’o | **llh**e- | | iXo | llhuN- | |
| Tindi | Lower | ima | **llh**uN- | **llh**o- | | | |
| Botlikh | | eN=C | **llh**e- | | e=b | llhun- | |
| Godoberi | | e=w=u, e=j=i | **llh**ee | llhee-, **llh**ee= | e=b=u | llhun- | |
| Karata | | hemol | **llh**olal | **llh**obottl, **llh**obol | hedol | hede- | hede- |
| Karata | Tukita | chu=w, chu=j | **llh**o- | **llh**o=C | chu=bi | **ss**un- | |
| Bagwalal | | he=C | llho- | **llh**o=C | he=C | he-llhi- | |
| **Tsezic:** | | | | | | | |
| Tsez | | show | **lh**u- | **lh**i | show | lhina | lhu- |
| Xvarshi | Inxokvari | hibo | **lh**u | lhi- | hibo | lhene | |
| Hinuq | | lhu | **lh**u- / l**h**i- | | **s**e | lhin- | |
| Hunzib | | suk’u | suk’u | sA- | shijo | **s**ëlë; **s**in- | **s**ëj |
| Bezhta | Tljadal | suk’o | **lh**o / **lh**oo- | | shizhö/shijö | lhini | |
| Lak | standard | cwu | **ssh**ii- | | **c**i | **ss**a- | |
| **Dargwa:** | | | | | | | |
| Dargi | standard | chi | chi | chi- | **s**e | se- | se- |
| Chiragh | Chiragh | cha | **ssh**-i- / **ssh**-e- | | **c**e | ci- | c- |
| Kubachi | Kubachi | cha | chi-di-l | chi- | **s**e | si-l-di-l | si-li- |
| Mehweb | Mehweb | chi-j-a | hi- | | si-j-a | se-lV- | |
| Sanzhi-Ic’ari | Ic’ari | cha | hi- | | **c**e | ce- | |
| Xaidaq | Xaidaq’ | cha | chi-li | chi-la | **c**i | **c**i-lli | **c**i-lla |
| **Lezgian:** | | | | | | | |
| Lezgi | standard | wuzh | ni | ni, ne- | wuch | kü | kü-/kwe- |
| Lezgi | Axty | fimi | ne- | | wish | chü | |
| Tabassaran | North | huzhu | **sh**i | | fi | fitta | Pl. fij- |
| Tabassaran | South | fuzh | **sh**-li | | fu | ftti | Pl. f-j-ir- |
| Aghul | Burschag | fush | na | she- | fi | firi | |
| Aghul | Richa | fish | na | | fi | fi-tt-i | |
| Aghul | Burkixan | fish | hana | | fi | firi | |
| Aghul | Fite | fizh | hina | | fi | [n.d.] | |
| Rutul | | hushi | hal | | **sh**iwi | hiji- | |
| Tsaxur | Mixik | hushii | **sh**awan | **sh**awun- | hizhoo | nishinne | nishiste |
| Kryz | | tij | an9a- | holi- | shi | chidzhi- | |
| Budukh | | tu | hanyr, hanyz | hanu | shi | hanyr, hanyz | hanu |
| Archi | | kwi | **llh**i | llha- | hani | hin- | |
| Udi | | **sh**u | **sh**ina | **sh**uva, **sh**i | hikä | hetin | hetu, hetaj |
| Khinalug | | kla | ksh- | ksh- | ja / jaaza | chin- | |

Table 2. Interrogative pronouns: 'who', 'what'

than nominative stems. This suggests that the nominative forms are separate intrusions into an older paradigm which survives in the oblique stems.

Root elements that occur systematically in oblique stems include the following (bold in Table 2).

*llh- (Gigineishvili 1977: 128), a fortis voiceless lateral fricative, chiefly in the oblique stem of 'who' (extending to the nominative in several Lezgian languages). This element displays regular correspondences in all Daghestanian branches: laterals in Avar-Andic and Tsezic and Archi, uvular or velar fricative elsewhere, depending on the language. (Velar reflexes have often become *sh*-before a front vowel: Lak oblique *sshii-,* Chirag Dargwa *ssh-i,* Tabassaran *shi*.) Though above I proposed that in the second person pronoun Nakh *h9* (voiceless pharyngeal fricative) stems from a Proto-ND uvular, in 'who' the same Nakh consonant may also be a regular reflex of *llh-. I have only one good cognate set for *llh-: 'water' (Nichols 2003: 263 #54): Chechen *xi*, obl. *xin-*; Avar *llhin*, Lak *sshin*, Archi *llhan*, etc. In this set Nakh has *x* rather than the *h9* found in the oblique stem of 'who'. Since lateral *ttl' regularly yields both *9* (voiced pharyngeal) and *q'* in Nakh (Nichols 2003 and unpublished later cognate sets), it is plausible that both 'water' and the oblique stem of 'who' could have regular reflexes of lateral *llh. Alternatively, perhaps the Nakh *h9-* corresponds to the *h-* found in several Lezgian and Dargwa oblique stems. Both correspondences are phonetically plausible, but with so few cognate sets it is impossible to know which (if either) is regular or what may have conditioned the phonetic outcome.

*st-, the only onset cluster in the protolanguage, found chiefly in the oblique stem of 'what' but also 'who' in Hunzib. The Daghestanian languages regularly merge *st, *st', *stt with *c, c', *cc or *ss (Nichols 2003). An example of a lexical cognate set is 'bile': Chechen *stim*, Avar *ccin*, Akhvakh *ssimi*, Hunzib *simi*, Lak *ssi*, Dargi *himi*, Lezgi *seb*, Tsakhur *cimis/simis*, Archi *ssam*. The stem of 'what' shows these reflexes in several languages.

If the original distribution of these two oblique stem formatives was *llh-in 'who' and *st- in 'what', there has been some extension and redistribution in several of the languages. The oblique stem of 'what' has been extended to the nominative in Dargwa (e.g. standard Dargi *se*, Chiragh *ce*), and a new nominative stem (shared by 'what' and 'who') has displaced the old oblique stem of 'what' in some Lezgian languages (e.g. Lezgi *wuzh* 'who', *wuch* 'what'; Aghul *fush/fish* 'who', *fi*, obl. *fi-* 'what').

Other elements can be detected but have a much more limited distribution and cannot be reconstructed beyond individual branches. Of the nominatives, Nakh *m-* in 'who' is unique to Nakh (it is discussed below). The postvocalic -*m-* of a few Andic languages (Andi *emi*, Chamalal dial. *im-*, Tindi *ima*, Karata *hemol*) and Lezgi *fimi*, all limited to the nominative of 'who', may be the same element. Dargwa *ch- in 'who' is chiefly Dargwa. The correspondence *f:h* in nominative 'who' and both nominative and oblique 'what' in Lezgian recurs and is regular, though the reconstruction is uncertain (Gigineishvili 1977: 122–124). In 'what', the Nakh correspondence Ingush *f/h*, Chechen *h*, Batsbi *v* is regular but I am unsure of the ancestral form so I cannot trace it beyond Nakh.

## 4.   Interrogative adverbs: 'when', 'where', etc.

Table 3 gives the forms, which show a good deal of variety together with occasional patterns of generalization. Nakh uses the interrogative base *m- in nearly all of these adverbs; this is evidently the same element as in the nominative stem of 'who'. The Nakh oblique stem of 'what', *ste:n-, appears, suffixed, in Chechen 'where' and Chechen, Ingush 'why', and in the nominative stem of Chechen 'what' and Batsbi 'why'. The elements suffixed to these bases vary across the forms and sometimes among the three Nakh languages.

   Several Daghestanian languages generalize one or another form across several adverbs, using it as an interrogative base. Avar has generalized ki-. Several Andic languages generalize a form like *(h)in- across the various adverbs, and some generalize a form in ch- to a few adverbs. Forms in ch- undergo some extension in individual Andic and Dargwa languages, perhaps spreading from the pan-Daghestanian *chVm 'how many' discussed just below. Only in Nakh and Dargwa is a generalized interrogative base found not only among the adverbs but also among the pronouns: Nakh m- and Dargwa ch- are both also found in 'who'.

   Postvocalic elements are also diverse. Postvocalic *-n (i.e. *VN, hVn-) occurs in a number of different adverbs, without any single identifiable function. (Alekseev 1985: 74 traces most of the Lezgian examples to a coordinating particle *-na.)

   All Daghestanian branches have a form like *chVm reflected regularly in 'how many' and occasionally other forms, possibly also including 'who' in Dargwa. Dargwa and Lezgian have a form with *mu- in 'when'. These are the only instances where a root can be associated with a particular adverbial meaning.

   Experience with Indo-European leads one to expect to find a single interrogative root like IE *kw- occuring in most interrogative forms, but this is not the case with ND. The variety of forms is considerable even within branches, and no single interrogative root can be reconstructed for Proto-ND. The clearest pattern to suggest itself is that in Nakh *m- expanded from some source to make a near-complete sweep across the adverbs and become a general interrogative base. The source form may have been 'when', as that has m in Dargwa and Lezgian (other than the initial *m, however, the forms in Dargwa, Lezgian, and Nakh are not cognate). It is not clear how the sweep happened; of the post-root elements in the Nakh adverbs, some may be cognate to case endings but their functions differ, and some are not recognizable case endings, so the interrogatives do not look like the remains of a disintegrated case paradigm. Dargwa *ch- made a less sweeping extension across the interrogatives but, like Nakh m, also extended to 'who', where it took root not only in the nominative but also in the oblique stem.

   I believe it is fairly common cross-linguistically for 'what' to extend to interrogative adverbs when expressions like English what time, what for oust former when, why. This seems to have happened occasionally in Daghestanian, e.g. Northern Akhvakh and Tukita Karata chu- 'what', chuNda 'when', but it has not been systematic. 'What' has not been involved at all in the sweep of m across interrogative adverbs in Nakh. Rather, it is 'who' that shares a generalized interrogative base, in Nakh (m-) and Dargwa (*ch-).

Blank = form not found.

| Branch, Language | Dialect | when | where (at) | where (to) | how much | how many | how | why |
|---|---|---|---|---|---|---|---|---|
| **Nakh:** | | | | | | | | |
| Ingush | | maca | mycha | mycha | mel | massa | myshta | hana; senna |
| Chechen | lowland | maca | michahw; stianga | micha; stianga | mial | massa | muuxa | hunda; sti-anna |
| Batsbi | | macaN | miche(h9) | michE, mich | melh | | moh9 | vuN |
| **Avar-Andic:** | | | | | | | | |
| Avar | standard | kida | kib | kibe | ki9an | chan | kin | shaj |
| Avar | Antsux | kída | kib | kix | ka9anzha=b | cham | | |
| Andi | | innal / inna- | inu-… | i=C=al | eriGa-/ eriGwa- | chom- | | |
| Akhvakh | northern | chuNda | hagittli | hagaje | husshtuda | chami | | |
| Chamalal | L. Gakv. | inna | ittl'aa | ilhal | itl'ol | shaaN | itl'(-qa) | ella~etila; ittl'a: |
| Chamalal | Gigatli | inna | inna-… | inuk'o | hint'o | chamu | | |
| Tindi | Lower | hindala | inila | inila | hiNtl'uhubla | chwamila | | |
| Botlikh | | inda | | | | | | |
| Godoberi | | indjaqiw | ijaqi | ijaqi | intto | chamu | intl'axxu | intl'avu |
| Karata | | hiNdal | hiNgol | hiNgol, hiNdir | chami, hiNshdo'o | chami, hiNshdo'o | hiNshdal | heNssol |
| Karata | Tukita | chuNda | | | chaNcil, chami | chaNcil, chami | | ssuna |
| Bagwalal | | hinda | hindi | hindi | chwam/chom | chwam/ chom | heshta | heshta, hedzhaa |
| **Tsezic:** | | | | | | | | |
| Tsez | | neti | na | naghor | dice | dice | didur; bilhe | lhinatlaaj; shina |
| Xvarshi | Inxokvari | ito | na | naghul | doccu | shomo | | |
| Hinuq | | nete | ni | nido | deche | somo | | |
| Hunzib | | hydë; hëdë | nijo | nijo | hidil | hidil | hinaa | surba |
| Bezhta | Tljadal | nito | naa | naa, naadaa | laso | laso | | |
| Lak | standard | ta | chuu | chun | cikssa | cimi | cukun | can, civan |
| **Dargwa:** | | | | | | | | |
| Dargi | standard | murt | chinab(a) | china | secad | chum | sen-sen, se- | |
| Chiragh | Chiragh | murt | kala | kala | cice | chum | | |
| Kubachi | Kubachi | muut | chi-na=b | chi-na | kwaasib, kwaashud | chum | sag-hunne | silij |
| Mehweb | Mehweb | murta | | kuda | | | | |
| Sanzhi-Ic'ari | Ic'ari | murti | china=b | china | kusa | kusa | kuuti | celij |
| Xaidaq | Xaidaq' | murt | kwacci / kacci | kwacci / kacci | chum-… ? | chum=Ca | cigni | cilizh |
| **Lezgian:** | | | | | | | | |
| Lezgi | standard | mus | hina | hiniz | hiq'wan | shumud | hik'(a) | wuchiz |
| Lezgi | Axty | mys | fini | finiz | fiq'adar | shmid | | |
| Tabassaran | North | ha9la | na7an | na7ana | hiq'an | hiq'an | | |
| Tabassaran | South | fila | na7an | na7ana(dzhi) | fuq'an | fuq'an | | |
| Aghul | Burschag | mus | nandi | nandi | fah9an | shimur | | |
| Aghul | Richa | mus | hindi | hanich | fidah9an | ximud | | |
| Aghul | Burkixan | mus | nandi | nanch | fidah9an | ximud | | |
| Aghul | Fite | mus | nant'a | najich | fidah9an | ximud | | |
| Rutul | | mys | hily | hile7 | shumu= | shumu= | | |
| Tsaxur | Mixik | mysa | njaGha | njaqa | nimenne | xule= | | |
| Kryz | | mys | hotä | hotan | häxkärä | häxkärä | | |
| Budukh | | mys | heje | heje | hekke | hekke | shyma | shimat'i, chun |
| Archi | | basa | danna | dashi | shumej= | shumej= | hanzhu-gur | daki |
| Udi | | he-vaXt' | maja | maja | heqqadar | heqqadar/heqqara | | |
| Khinalug | | t'aga, mys | del | ttak'ol | chutt'on | chutt'on | chim | china, cchimä |

Table 3. Interrogative adverbs: 'when', 'where', etc.

## 5.  Similar developments in nouns

Some nouns in ND have what I call *overt inherent gender*, where the noun itself carries a prefix marking its own gender (as with Bantu prefixes or Spanish *-o* and *-a* endings). Usually the prefix is the same as the main or most transparent allomorph of the gender agreement prefix, but sometimes a prefix undergoes distant nasal assimilation to a root consonant. And occasionally there is spontaneous nasalization not conditioned by the phonology of the word. Table 4 shows the words for 'sun' and 'moon'. Both words belong to the B gender class (class 3) in all languages having three or more genders, and the initial *b is an overt inherent gender prefix for that gender. (The Lezgi and Xinalug cognates for 'sun', and the Lak oblique stem for 'moon', show that the word also occurs without the initial gender marker, which proves that the initial is a separate morpheme that was not present in all forms of the words in Proto-ND or Pre-Proto-ND.)[10] In each cognate set one or two subgroups change the initial to a homorganic nasal, though neither word contains (or contained) a nasal that might trigger nasal assimilation; and it is different subgroups in each word. This indicates that non-phonological replacement of /b/ by /m/ (or perhaps it is replacement of just the gender prefix *b=*) is a sporadic change that occurs occasionally in one or another word in the six or more millennia of Nakh-Daghestanian history.

| | 'sun' | | 'moon' | |
|---|---|---|---|---|
| Nakh: | | | | |
| Ingush | *maalx* | | **b**utt | **b**ett- |
| Batsbi | *matx* | | **b**utt | **b**att- |
| Avar: standard | **b**aq' | | *mocc'* | *mocc'rol* |
| Andic: Andi | *milhi* | | **b**orcc'i | |
| Tsezic: Hunzib | **b**oq | **b**yq- | **b**oco | **b**yc- |
| Inxokwari | **b**yq | | **b**ucu | |
| Lak | **b**argh | **b**urgh- | **b**arz | zur- |
| Dargwa: Dargi | **b**erh9i | | **b**az | |
| Xaidaq' | **b**ari | | **b**ac | |
| Lezgian: | | | | |
| Lezgi (standard) | **r**agh | **r**aq- | **w**arz | **w**acra |
| Budux | **v**iragh | | **v**əz | |
| Archi | **b**arq | | **b**ac | **b**ocro |
| Udi | **b**egh9 | | | |
| Xinalug | ynq' | | **v**ac' | |

Table 4. Nakh-Daghestanian cognate sets of B gender with overt inherent gender marked by their initial consonants (bold). Oblique stem shown only if different from the nominative. Selected representatives of each branch. Words with nasal initial are italicized.

---

10.     The Nakh languages retain a handful of words that have overt inherent gender initials in the nominative but not in the oblique cases. For Nakh see Nichols 2007: 1182–4, 2011: 147–9. Two examples are Ingush *joq'*, Chechen *juq'*, obl. *ouq'ar-* 'ashes' and *jexk*, oblique *axkara-* 'comb' (both J gender).

# 6. Discussion

Thus we see that most tokens of initial *m-* in Nakh-Daghestanian pronominals are secondary: all those in the second person singular, all or most of those in the Nakh interrogatives, and possibly the occasional postvocalic one in nominative 'who' in Andic and Lezgi. The reasons for the extensions of *m* are different: phonetic or phonological in the Daghestanian second person pronouns, morphological in the Nakh interrogatives, unknown for the postvocalic ones. The source of the interrogative *m-* is not entirely clear. Only the word 'when' has *m-* in different branches (Nakh, Dargwa, Lezgian), but as only the *m-* is cognate across the three it is not possible to reconstruct a whole Proto-ND word for 'when'.

Though secondary, *m-* is quite stable. I know of no sound changes in which it is lost in ND, and among the forms reviewed here there are no cases of *m-* being ousted in morphological change. Recall from §2 above that in Avar-Andic and Tsezic the reflex of Proto-ND *gh(w) was ousted or renewed in the oblique forms by *d-* imported from the first person paradigm; but in the oblique forms that reflex was almost certainly something other than [m].

*m-* also displays a modest tendency to expand further from the nominative to the ergative in Andic and Tsezic. Sometimes, as in the second person form of Andi, Godoberi, and most Tsezic languages, there is full nominative-ergative syncretism (e.g. Hinuq *me* 'you', nom.=erg.); sometimes the nominative is taken over as the ergative stem and further inflected. This system seems quite stable, and *m-* has not spread further into the oblique cases.

In Nakh, *m-* is expanding from the interrogatives to other pronominals: both *massa* 'how many' and *mel* 'how much' figure in expressions for 'all' and 'every' which are displacing the older pan-Nakh form, e.g. Ingush *massa xaana* 'always', lit. 'how(ever) many times', *shie mel d.ar* (REFLEXIVE – how much – be-NOMINALIZER) 'all of them' (lit. 'however many themselves were').

A recurrent tendency throughout the family is to replace the nominative of 'who' and/or 'what' with an etymologically unrelated neologism, retaining the inherited oblique paradigm. How and why would a neologism have entered in only the nominative? One possibility is that the neologism originated as an emphatic form used only as predicate nominal. A familiar modern example of such a distribution is the circum-Baltic expression German *was für*, Russian *chto za*, etc. This phrase figures chiefly as predicate in questions meaning 'what kind of X is Y?'. It evidently started out as an emphatic or unlimited-set form meaning 'what on earth kind of…', 'who in the world…' and gradually displaces its original near-synonyms to more restricted-set meanings like 'which one'. The semantics of 'who' and 'what' is not the same as 'what kind', and the morphological fate of the ND neologisms is not to displace another lexeme but to replace the inherited nominative in an existing paradigm.

Another possible entry point is suggested by modern Nakh syntax, where 'who' and 'what' have full case paradigms but are in fact not often used in the oblique cases. Rather than use a non-nominative interrogative, natural speech

much more often uses a nominalized clefting construction that keeps the focus
of interrogation in the nominative, e.g. Ingush:

(1)   Hwan / hwanuo yz deadar?
      who.ERG / who-ERG it D.do-D.PAST NONWITNESSED (D = a gender prefix)
      Who did it?

(2)   Malu=u yz dear?
      who=is.V it D.do-PARTICIPLE.NOMINALIZER  (D, V = gender prefixes)
      Who did it? (Lit. Who is it that did it?)

Examples like (2) considerably predominate in natural texts. In (2) the interroga-
tive is semantic predicate and syntactic predicate nominal with copula.

   Thus there are some natural avenues for renewal of just the nominative in
an interrogative paradigm. There will probably never be enough morphosyn-
tactic information to reconstruct a specific source for any of the innovative ND
nominatives.

   There are some additional consonants with appreciable propensity to ex-
pand, all of them phonetically strident obstruents: the oblique stems of 'who',
*llh- and its various fricative reflexes; the oblique stem of 'what', *st- and its
reflexes st-, c-, s-; *ch- in 'how many', which becomes a generalized interroga-
tive base in Dargwa. Proto-ND *d- (which may have been more strident than this
spelling would suggest: see again §2) and its reflexes are quite stable in the first
person singular pronoun and, if the suggestion of §2 above is correct, may have
spread to the second person singular as well. These developments indicate that a
coronal obstruent has good stability and some tendency to spread, and may also
be an attractor state.

   The two consonants even form a minimal paradigm in the Avar-Andic and
Tsezic nominatives 1sg. *dun*, 2sg. *mun* (cited from standard Avar), reminiscent
of the widespread Eurasian pronouns with *m* in 1sg and some coronal obstruent
in 2sg. – the consonant mini-paradigm is the same, but the values are reversed,
with *m* marking first person in northern Eurasia but second in Daghestan.
Now, the sociolinguistics of the Avar-Andic-Tsezic area is distinctive in the
Caucasus.[11] For at least three millennia, from the rise of the Sarir kingdom (later
known as the Avar Khanate after its conversion to Islam) to the mid-19th century
when Russia defeated the Avars and their allies, there was a stable center of
military and economic power in the east Caucasus foothills, in and near the Avar
capital Khunzakh. Language spreads from lowlands to highlands accompanied
this power configuration. In historical times an Avar variety was the language
of command in the military, and highlanders in the Avar sphere of influence
needed to know Avar because the markets and winter pastures were in the Avar
lowlands. The linguistic influence continues today, as the rural economy is

---

11.    This analysis is summarized from Nichols 2005, in press.

still vertically based and in the early Soviet years Avar was made the official ethnicity and language of elementary school instruction among small ethnicities in the Avar sphere and continues to be an important lingua franca.

This situation appears to be long-standing. The Andic family appears to be less than 3000 years old, and its distribution along the middle Andi Koisu probably results from an uphill spread of Proto-Avar-Andic before the more recent spread of specifically Avar. The Tsezic range in the upper Andi Koisu may result from the spread of a still earlier Proto-Avar-Andic-Tsezic. The inexorable uphill spread of the lowland language proceeded not just by gradual upward movement of isoglosses but also by sporadic leapfrogging of Avar into the upper highlands. Meanwhile, the highland towns were politically and ethnically autonomous, essentially city-states, and using Avar or indeed even shifting to Avar did not change this status.[12] The city-state, rather than the Avar capital, was the focus of ethnic identity (in addition to the higher-level identities of Muslim and mountaineer). The complexity and multiplicity of linguistic interactions, the repeated spreads of different languages from the foothills, the lack of any attempt to actually impose Avar as everyday language, the stability of ethnic identity, and the independence of ethnic identity from language all recall the complex and multi-centered linguistic situation that must have characterized the eastern and Mongol steppes and eastern Central Asia from the first spread of the Neolithic to the eastern steppe, eastern Central Asia, Mongolia, and Tibet to the time of Genghis Khan (and longer in the Mongol hinterlands).[13] If, as I have argued, the spread of a common pronominal canon in eastern Eurasia results from that sociolinguistics enhancing the prospects of an attractor state, something similar may have happened in the Avar-Andic-Tsezic area. Meanwhile, the sweep of *m* across the Nakh interrogatives and the Avar-Andic-Tsezic second person pronouns, and the sporadic changes of gender-marking *b to *m* in 'sun' and 'moon' (and other lexemes), show that innovations of this particular attractor state are sporadic and not frequent, appearing once or twice in a handful of cognate sets over the six millennia or so of Nakh-Daghestanian prehistory. But once in place they are durable enough to be conspicuous in the modern languages.[14]

---

12. Aglarov 1988a uses Russian the term *obschina*, which is often translated 'commune', capturing aspects of their internal political and economic structure. He also describes Daghestanian towns as similar to Greek city-states. For discussions of language identity and language shift, I believe the comparison to city-states is more appropriate.

13. For more on the language situation at the Eurasian steppe periphery see Janhunen 2003a, Johanson 1998, Nichols in press. The post-Neolithic sociolinguistics of the eastern steppe accounts for the closely resemblant personal pronoun systems of Turkic, Tungusic, and Mongolic but not for the less closely resemblant and geographically more distant Uralic, Indo-European, and Kartvelian. These display the same attractor state (first person *m*, second person coronal), but I leave open the question of whether they bear a historical (as opposed to typological) connection to each other and to the eastern steppe. Independent development is the most parsimonious assumption for the individual systems, but it does not explain why there is a cluster of such systems in northern Eurasia when they are rare in the rest of the world (Nichols & Peterson 2005).

# References

Aglarov, M. A. 1988: *Sel'skaja obschina v Nagornom Dagestane v XVII-nachale XIX v.* Moscow: Nauka.

Alekseev, M. E. 1985: *Voprosy sravnitel'no-istoricheskoj grammatiki lezginskix jazykov. Morfologija. Sintaksis.* Moscow: Nauka.

—— 1988: *Sravnitel'no-istoricheskaja morfologija avaro-andijskix jazykov.* Moscow: Nauka.

Carlson, Barry F. and John H. Esling 2003: Phonetics and physiology of the historical shift of uvulars to pharygeals in Nuuchahnulth (Nootka). *Journal of the International Phonetic Association* 33(2):183–193.

Creissels, Denis 2009: Uncommon patterns of core term marking and case terminology. *Lingua* 119(3): 445–459.

Gigineishvili, B. K. 1977: *Sravnitel'naja fonetika dagestanskix jazykov.* Tbilisi: Tbilisi University Press.

Greenberg, Joseph H. 1960: An Afro-Asiatic pattern of gender and number agreement. *Journal of the American Oriental Society* 80: 317–321.

Gudava, T. E. 1959: *Sravnitel'nyj analiz glagol'nyx osnov v avarskom i andijskix jazykax.* Maxachkala: IIJaL Dag. Fil. AN SSSR.

Haas, Mary R. 1958: Algonkian-Ritwan: The end of a controversy. *International Journal of American Linguistics* 24: 159–173.

Haspelmath, Martin 1993: *A Grammar of Lezgian.* Berlin: Mouton de Gruyter.

Holisky, Dee Ann and Rusudan Gagua 1994: Tsova-Tush (Batsbi). In Rieks Smeets, ed., *The Indigenous Languages of the Caucasus,* vol. 4. Delmar, NY: Caravan. 147–212.

Janhunen, Juha 1981: Uralilaisen kantakielen sanastosta. *Journal de la Société Finno-ougrienne* 77: 219–274.

—— 1982: On the structure of Proto-Uralic. *Finnisch-Ugrische Forschungen* 44: 23–42.

—— 1996: Prolegomena to a comparative analysis of Mongolic and Tungusic. In Giovanni Stary, ed., *Proceedings of the 38th Permanent International Altaistic Conference.* Wiesbaden: Harrasowitz. 209–218.

—— 2001: Indo-Uralic and Ural-Altaic: On the diachronic implications of areal typology. In Christian Carpelan, Asko Parpola and Petteri Koskikallio, eds., *Early contacts between Uralic and Indo-European: Linguistic and archaeological considerations*: 207–220.

—— 2003a: Proto-Mongolic. In Juha Janhunen, ed., *The Mongolic Languages.* London: Routledge. 1–29.

—— 2003b: Para-Mongolic. In Juha Janhunen, ed., *The Mongolic Languages.* London: Routledge. 391–402.

—— [in press]: Non-borrowed non-cognate parallels in bound morphology: Aspects of the phenomenon of shared drift with Eurasian examples. In Lars Johanson and Martine Robbeets, eds., *The Origins of Bound Morphology.* Leiden: Brill.

Johanson, Lars 1998: The history of Turkic. In Lars Johanson and Éva Ágnes Csató, eds., *The Turkic Languages*. London: Routledge. 81–125.

Kibrik, A. E. and S. V. Kodzasov 1990: *Sopostavitel'noe izuchenie dagestanskix jazykov: Imja. Fonetika*. Moscow: Moscow University.

Kibrik, Aleksandr E. (ed.) 2001: *Bagvalinskij jazyk*. Moscow: IMLI RAN "Nasledie".

Kibrik, Alexandr E., Sergej G. Tatevosov and Alexander Eulenberg 1996: *Godoberi*. Munich: Lincom.

Magometov, A. A. 1963: *Kubachinskij jazyk*. Tbilisi: Georgian Academy of Sciences.

—— 1982: *Megebskij dialekt darginskogo jazyka*. Tbilisi: Mecniereba.

Meillet, Antoine 1925: *La Méthode comparative en linguistique historique*. Paris: Champion.

Newman, Paul 1980: *The Classification of Chadic within Afroasiatic*. Leiden: Universitaire Pers.

Nichols, Johanna 2003: The Nakh-Daghestanian consonant correspondences. In Dee Ann Holisky and Kevin Tuite, eds., *Current Trends in Caucasian, East European, and Inner Asian Linguistics: Papers in Honor of Howard I. Aronson*. Philadelphia: John Benjamins. 207–251.

—— 2005: The origin of the Chechen and Ingush: A study in alpine linguistic and ethnic geography. *Anthropological Linguistics* 46: 129–155.

—— 2007: Chechen morphology (with notes on Ingush). In Alan S. Kaye, ed., *Morphologies of Asia and Africa (including the Caucasus)*. Winona Lake, IN: Eisenbrauns. 1161–1180.

—— 2011: *Ingush Grammar*. Berkeley-Los Angeles: University of California Press.

—— [in press]: Selection for *m : T* pronominals in Eurasia. In Lars Johanson and Martine Robbeets, eds., *The Origins of Bound Morphology*. Leiden: Brill.

—— [submitted]: A direct/inverse subsystem in Ingush deictic prefixes.

Nichols, Johanna and David A. Peterson 1996: The Amerind personal pronouns. *Language* 72: 336–371.

—— 2005: Personal pronouns: M-T and N-M patterns. In Martin Haspelmath, Matthew Dryer, Bernard Comrie and David Gil, eds., *World Atlas of Language Structures*. Oxford: Oxford University Press. 544–551.

—— 2011: Personal pronouns. Matthew S. Dryer and Martin Haspelmath, eds., *The World Atlas of Language Structures Online*. Munich: Max Planck Digital Library, chapters 136–137. Available online: http://wals.info/chapter/136 (Accessed Nov. 4, 2011)

Nichols, Johanna and Arbi Vagapov 2004: *Chechen-English and English-Chechen Dictionary. / Noxchiin-ingals, ingals-noxchiin deshnizhaina*. London: Taylor and Francis.

Nikolayev, S. L. and S. A. Starostin 1994: *A North Caucasian Etymological Dictionary*. Moscow: Asterisk. (Reprint, 2007: Ann Arbor: Caravan.)

Sapir, Edward 1913: Wiyot and Yurok, Algonkin languages of California. *American Anthropologist* 15: 617–646.

Schrijver, Peter 2009: Why Indo-Europeanists should know more about East Caucasian. MS, Universiteit Utrecht.

Schulze, Wolfgang 1999: The diachrony of personal pronouns in East Caucasian. In Helma van den Berg, ed., *Studies in Caucasian Linguistics*. Leiden: University of Leiden. 95–111.

Sumbatova, Nina R. and Rasul O. Mutalov 2003: *A Grammar of Icari Dargwa*. Munich: Lincom.

Temirbulatova, S. M. 2004: *Xajdakskij dialekt darginskogo jazyka*. Maxachkala: Russian Academy of Sciences, Daghestan Nauchnyj centr.