

**On the substratal heritage of the Central Asian Sprachbund**  
*(К субстратальному наследию Центральноазиатского языкового союза)*  
By Joy I. Edelman  
Published in 1980 (in *Вопросы языкознания*)  
Translated by Clinton Parker, 2023

The intensive comparative research during the past years into the languages of the Indo-Iranian family has allowed us not only to trace the fundamental tendencies of spontaneous development of their source languages, but also to identify a series of interesting ‘exceptions’, ‘deviations’, ‘anomalies’, not motivated, as it were, by the rule-based evolution of these languages. A part of these anomalies has arisen, apparently, due to idiosyncratic ‘mutations’ of the system (cf. the reconstruction of the deictic systems of many Indo-Iranian languages with triple deictic systems to previous binary systems through the use of etymologically different components; the emergence of various types of ergative or ergative-like constructions; and, connected with the latter, the ‘inversion’ of the copula in Pamir languages as a result of its contamination with enclitic pronouns; etc.). A part can be explained through the convergence and borrowing of elements from other languages over the course of lengthy contacts (for example, retroflex consonants in Baluchi, which were assimilated together with lexical items and *izofat* constructions in languages which had contact with Persian and Tajiki; etc.). In the end, a part of the transformations in the group of languages of a single area (though not necessarily closely related languages) can be caused through substratal influence. The latter can serve as a catalyst for tendencies which potentially existed in the proto-language, but which were not realized in other languages, though it can evoke changes and illogical developments in the original system.

The analysis of such type of ‘anomalies’ can give information both on the history of the relevant languages and on the structure of the substrate language – with respect to its ‘superficial’ (formal) structural characteristics and the phenomena which reflect its unified content-typological makeup.

In the special literature there has already been indications of substratal features on a variety of levels: in phonetics, in word formation, in inflectional models, in syntax, in the lexicon, and also in general typological characteristics.<sup>1</sup> In this article, we summarize the anomalies which can be attributed to substratal influences in genetically homogenous languages of the same area.

This area includes the regions of the Hindu Kush, the Pamirs, spurs of the Karakoram, and part of the Himalayas and includes languages of various genetic belonging. Such languages are Eastern Iranian (e.g. Pashto, Pamir, Munji); Western Iranian (Baluchi); and two relict Iranian languages (Parachi and Ormuri, which display both Eastern and Western features). There are also some Indo-Aryan languages throughout this region: Domaaki, the so-called dialects of the Pahari people (of Nepal), Punjabi, and also due to some features the Lahnda languages and Sindhi; Dardic languages also belong fully to this area. The Nuristani languages (traditionally Kafir languages) also belong to this group, which, alongside Indo-Aryan and Iranian, make up the third branch of the Iranian language family.<sup>2</sup> The obvious non-Indo-European language here

---

<sup>1</sup> Abaev, V. I. 1956; 1970; 1978. V.S. Vorobev-Desyatovskij. 1956; Edelman 1968; 1976

<sup>2</sup> On the position of these languages within Indo-Iranian, see: Morgenstierne 1973; 1974; Buddruss 1977.

is the language isolate Burushaski, along with some Sino-Tibetan (Himalayan) languages and individual Dravidian languages (Brahui and possibly certain dialects in Nepal). A series of multi-level structural parallelisms allow us to apply the notion of a Sprachbund to these languages, which is called in the literature Central Asian or Himalayan Sprachbund.<sup>3</sup>

Of the purely formal characteristics of the languages of this area which can be explained through substratal influence, we can mention the following.

There is a certain generality in their phonological structure: complexity in their consonantal system and relative simplicity in their vowel system; the areal fixation of certain phonological series – in particular, retroflex fricatives (and in an especially narrow area, also retroflex affricates). This is in contrast to the majority of Indo-Aryan languages, where retroflex series of consonants are only found in stops,<sup>4</sup> but is in correspondence with the phonological inventory of non-Indo-European languages of the region, particularly Burushaski.<sup>5</sup>

There is a specific areal tendency toward certain word-formational models, which points, first and foremost, toward a substratal origin. Thus, the model of forming the second-person plural pronoun is not from the old stem of the full or enclitic pronoun, but rather based on the scheme *\*tu + sma* (for the vast majority of Indo-Aryan languages, including a part of Dardic languages: Wotapuri, Tirahi, Torwali, Bashkarik, Mayan, Palula, Shina, Kashmiri) or *\*ta + hma* (for some Iranian languages: Pashto, Ormuri, Shughni-Rushani, Yazghulami, and Ishkashimi). This model could have had as a prototype the substrate model for forming the pronoun ‘you (pl.)’, which is formed from ‘you (sg.)’ attached to a plural marker<sup>6</sup> (cf. such a formation in Dravidian languages, Himalayan languages, and apparently in early Burushaski).

It is likely that a substratal schema is the explanation for the use of a numeral system in which the numbers 11-19 are built on the model *ten + two = 12* (unlike the typical Indo-Iranian system of *two + twelve*) in the languages of the area (in addition to Wakhi and a series of Dardic languages, where this order is connected to the analogous system in Burushaski (V.M. Beskrovnyj 1928), we can add the languages of the Shughni-Rushani group and Yazghulami). Substratal influence may also be the explanation for the vigesimal system of numerals above 30 in the majority of languages of the area, as well as in languages in more southern regions of Hindustan (cf. the vigesimal system in neighboring non-Indo-European languages, on the one hand, and the decimal system in late Indo-European and the fact that Indo-European languages are not native to the area, on the other).

We can also note here a series of special word-formation models. Thus, the Shughni expanded form of ‘two’ *điyûn* (cf. the primary form *đu*) is a form of the numeral and a nominal plural marker and cannot be linked to the history of Iranian languages (in old Iranian and Aryan languages *two* was included in the paradigm of dual number). As this numeral is used not only

---

<sup>3</sup> V.N. Toporov. 1965. See also the works indicated in fn. 1.

<sup>4</sup> Т. Я. Елизаренкова. *Исследования по диахронической фонологии индоарийских языков*. 1974: 278-280.

<sup>5</sup> И. И. Зарубин. *Вершское наречие канджитского языка*. 1927; Lorimer. *The Burushaski Language*. 1935. More information on Burushaski can be found in these sources.

<sup>6</sup> Morgenstierne 1935; V.S. Borob'ev-Desyatovskij 1956; Edelman 1971. For arguments against the borrowing of these pronouns by the Iranian languages from Indo-Aryan, see: F.B. J. Kuiper 1975; 1976.

in counting people, but also in counting objects and in abstract counting, it is similar to the classical Persian form *duvān* (in *har duvān* ‘a group of two’), but is not analogous to it in usage. With regard to its function and to its place in the system it is comparable to Burushaski *āltō* ‘two’ – the abstract form of the numeral (i.e. not used with nouns), which is unique among this series of numerals in that it is fixed with the nominal plural marker *-o*.

In the formation of nouns, consider, for instance, Wakhi *žəmák* ‘moon’; *ž(ə)yir* ‘sun’ with the fused element *žə-* ‘my’ (cf. pronominal proclitics in Burushaski, which are also used with nouns)<sup>7</sup>; *brin* ‘knee’; *brət* ‘elbow’, where *b-* < \**dba-* < \**dva* ‘two’<sup>8</sup> (cf. the analogous structure for the names of paired body parts in Burushaski with the prefixal *-lt-* < *āltō* ‘two’). Similar parallelisms are observed in word formation, function words and elements, and in the syntax of Shina, Khowar, and Dumaki, on the one hand, and Burushaski, on the other.<sup>9</sup>

In the lexicon we find certain semantic shifts (sometimes accompanied by phonetic irregularities): cf. Wakhi *yumj* ‘flour’ < \**āmači-* ‘raw’ and Burushaski *duyōun* ‘flour’ ~ *duyūi* ‘raw’. It is possible that we can add to the list of such shifts the case in several Pamir languages of the word (and Tajik dialects of Badakhshan) the words *dawōm*, *dawom*, *dawūm* ‘beginning’ as a nominal component in the complex verb meaning ‘start (tr./intr.)’, which is a contamination of the Arabic borrowing *dawām* through Tajiki *davom* ‘continuation’ with a similar word that formerly existed in Pamir languages meaning ‘start’. Such a word could have been a substrate word connected etymologically to Burushaski *dūn-* ‘hold’, especially considering that this verb in Wershikwar usually appears in the participial form in *-m* or in the absolutive *dohon*. Considering the absence in the Pamir languages of the phoneme /h/, such a form could have been inherited from the substrate language as \**dowom* (with *w* to resolve the hiatus; whereas a borrowing from Tajiki would have preserved the phoneme *v*) and the contaminated form *davom* (Arabic)...?

We can also observe here a noticeable layer of areal lexicon which does not always have a clear source, some of which can apparently be of substratal origin. Of course, samples of substratal lexicon should be looked for in toponyms, which are far from only having only Indo-Aryan origins.

The similarities discussed above have a fundamentally formal character. However, in terms of a possible substratal influence, we should also find interesting the features of the languages of this area which allow us to make specific conclusions of the substantive side of linguistic forms and reveal elements of that which are generally called ‘speech’ or ‘linguistic thought’, which stand behind the so-called ‘hidden categories – that is, elements which are often preserved upon the transfer of an ethnic group from one language to another. Precisely these elements, connected with the substantive side of the language, allow us, as it seems, to trace a certain system within the numerous structural anomalies of the languages of the area. If we accept Abaev’s definition of a substrate as the aggregate of the regular ‘mistakes’ of the sphere which does not get transferred onto the other language, then . . .

---

<sup>7</sup> Steblin-Kaminskij 1970

<sup>8</sup> Steblin-Kamenskij 1979.

<sup>9</sup> Lorimer 1937; 1968; 1939.

It should be noted beforehand that the geographic distribution of the elements in question, as well as the elements of a formal nature, are not homogenous (and have not been uniformly studied). A part of these elements may be common to two or more contiguous languages (though not necessarily closely related to one another)<sup>10</sup>; other features are scattered amongst individual languages. Because it is not possible to examine all of them in detail in this article, we limit ourselves to a few of the ones which, in our view, are the most characteristic.

In certain Eastern Iranian languages, in particular Pashto, the languages of the Shughni-Rushani group, Yazghulami, Munji, and also in the neighboring Baluchi and Parachi, we see a noteworthy anomaly with regard to the distribution of verbs into classes of transitive and intransitive. Some verbs with intransitive meaning (i.e. which do not take a direct object) are treated as transitive: either they are conjugated in the same way as transitive verbs or they call for ergative-like clausal constructions. These are verbs of the type: *sigh, sneeze, cough, laugh, smile, cry, feel nauseous, die, to be greedy, to be proud, to grow, become, jump, run, approach, sit, wander, dance, play, swim, to escape, to bark, meow, howl, crow, bleat*, etc., as well as verbs of directed motion – in other words, verbs indicating the action of an animate subject. The most extensive full list of such verbs for the Iranian languages is for Pashto (35 lexemes).<sup>11</sup> Somewhat lower is the number observed in Munji, for which the following are recorded: *laugh, cry, cough, become nauseous, agree, bark, defecate, stand, fall (of precipitation), knock, quarrel*.<sup>12</sup> The fact that we get the verb *to fall* (of precipitation) in this group is explained by the fact that a series of natural phenomena (precipitation, rainbows, movement of the skies, thunder, etc.) were previously understood to be the manifestation of an animate origin, which was then reflected in the language. Instances in which these verbs are treated as transitive are found in a number of Pamir languages: Yazghulami (examples); Rushani (examples); Bartangi (examples); Shughni (examples); etc.<sup>13</sup> Verbs of this type in Parachi are *sit, die, cry* (Morgenstierne 1929: 96); in Baluchi they are *laugh, run, sneeze, bark*, etc. (Frolova 1960:48). The same anomalous verbs are found in Dardic languages (Kashmiri, Shumashti, and dialects of Pashai) and in Indo-Aryan proper (Hindi, Assamese, Marathi, and Parya). Undoubtedly, more comprehensive research into these languages will fill out the list of verbs of this group and the list of languages which show this ‘anomaly’.

Intriguing are analogous deviations in the formulation of verbs of the same semantic group in bygone Iranian languages of the region – Sogdian (*approach, sit, stand, dawn*; etc.)<sup>14</sup> and Saka (*live, reach Nirvana, arrive, ride (a horse), cry, stray / deal with failure, become / be, grow, sleep, stay*). At that time, S. Konow identified the group of verbs in Saka and characterized them as “those intransitive verbs which have a heavily active meaning.”<sup>15</sup> Unfortunately, his formulation was not adequately valued and even received sharp criticism regarding the classification of these verbs as ‘exceptions’ which had taken on the formal distribution of the transitive type of conjugation.<sup>16</sup> However, the semantic characteristics of these verbs and a comparison of them with analogous verbs in other languages of the area supports the conclusions

---

<sup>10</sup> Cf. the common substratal features of Caucasian languages, Ossetic, and Armenian (V. I. Abaev).

<sup>11</sup> Avoryankov 1960.

<sup>12</sup> Sokolova 1973: 76, 100, 101, etc.

<sup>13</sup> See Fayzov 1966: 40; 155.; Karamkhudoev 1973: 153; Shughni examples were given kindly by Karamshoev.

<sup>14</sup> Gershevitch 1954: 130.

<sup>15</sup> S. Konow (1949: 50).

<sup>16</sup> R.E. Emmerick 1968:221.

of S. Konow (an ‘inverse exception’ is the verb *bud-* ‘be conscious of; know; understand’, at least seemingly: *verba sentiendi*, a group to which *bud-* can be classified, are treated in many languages either as intransitive or as a group on their own).

Noteworthy is the presence of a similar deviation in Burushaski (such verbs are ‘become intoxicated’, ‘go to bed’, ‘arrive’, and ‘laugh’).<sup>17</sup>

Therefore, in the languages of this area we observe a very specific feature: the link of a certain group of verbs with active meanings (i.e. the action of an animate being) to transitive verbs. In each of these languages, this group of verbs forms an exception to the general morphological and syntactic characteristics which distinguish (in)transitivity, though the uniformity of the semantics and consistency of these verbs from language to language would lead us to believe that this phenomenon is not accidental. It may point to a certain semantic principle along which the division of the verbal lexicon is divided, a principle by which verbs are divided based not on transitivity/intransitivity, but rather animacy/inanimacy of the subject. (A similar phenomenon is observed in Kurdish by I. I. Zuckermann, though here it is not accompanied by further anomalies which will be discussed below.)

In some languages of the area, it is characteristic for there to be two different copular lexemes for the third-person singular – one for animate subjects and one for inanimate subjects. This is seen in some Dardic languages (Pashai, Shumashti, Khowar, and Kalasha) and in an Iranian language (Parachi).<sup>18</sup> The lexical opposition of animacy and inanimacy in the copula (cf. Kashala pres.-singular *asou* ~ *šiu*, pl. *asan* ~ *šian*; past tense *asis* ~ *ašis*, *asini* ~ *ašini*; Parachi (*h*)*a* ~ *sī*; etc.) does not ensue from the Proto-Indo-Aryan (or Indo-Iranian?) system (even though both lexemes are etymologically native). Nevertheless, the special form of the copula in nouns of the Class IV (historically inanimate) is observed in Burushaski.

In some languages of this region, for instance in Dardic – Kashmiri, Shumashti, etc. – there is a series of verbs which indicate a quality, of the type ‘be red’, ‘be big’, ‘be bitter’, ‘be hot’, etc: e.g. Kashmiri *wazal-* ‘be red’; *pōṭh-*, *vyath* ‘be greasy’; *g<sup>o</sup>ab-* ‘be heavy’; *lōk-* ‘be small’; *čhat-* ‘be white’; *tat-* ‘be hot’; *ṭyaṭh* ‘be bitter’; *thad-* ‘be tall’; etc. Given that in Proto-Indo-Aryan there was a rich system of adjectives as a separate part of speech with a separate inflectional system, and given that in many Indo-Aryan languages adjectives have a tendency to link with nouns, the emergence in such a narrow geographical area of such verbs as predicates can hardly be attributed to the natural development of Indo-Aryan adjectives. Rather, this is an exception which can be compared to an analogous group of verbs in (Himalayish = Tibeto-Burman?) languages and in Burushaski (in Dravidian they appear at a later stage).

In certain languages we find a second lexeme for the first-person plural pronoun which opposes the other based on inclusivity/exclusivity. In Iranian languages – Yazghulami and Balochi – this opposition appears to be between inclusive/general: cf. Yz. *az-tow* ‘me and you’ (lit. *I-you*) ~ *mox* ‘we (general)’; Baluchi *mašmā* ‘me and you’ or ‘me and you (pl)’ (lit. *I-you* or *we-you*) ~ *amma* ‘we (general)’.<sup>19</sup> A similar phenomenon is also found in certain Indo-Aryan languages.

---

<sup>17</sup> Lorimer. *The Burushaski language*. p. 199.

<sup>18</sup> IIFL III, 1 (1967): 100, 176, 229, 274; Morgenstierne 1947:25; 1965:219-221; *Notes on Shumashti*, p. 225

<sup>19</sup> Sokolov 1956: 69.

In non-Indo-European languages of the region – Himalayish, Dravidian (and also in Munda) – the opposition between inclusive/exclusive is almost universal.<sup>20</sup> The Proto-Indo-Aryan language did not have such an opposition based on the participation or lack of participation of the interlocutor in a given situation, and the spontaneous development of the system could have hardly developed such an opposition naturally.

In some languages of the region, a special way of expressing the opposition between alienable/inalienable possession has been formed. It is implicit in the majority of Indo-Iranian languages – in the typical usage of terms of inalienable possession with a qualifier (pronoun or enclitic), whereas the rest of the nouns in a given language take such a qualifier optionally, depending on the meaning of the utterance. In some languages of the region this opposition is even expressed syntactically – through the order of elements in the prepositional phrase.

Thus, in a portion of Eastern Iranian languages – in the majority of Pamir languages, including Munji – possessive combinations with a preposition are typically built on the sequence *preposition – modifier – modifier* (of the type *ən ni kūd* ‘at my house’). In Pashto and in Sarikoli (the only one of the Shughni-Rushani group), the preposition comes in the middle of this combination (Psht. *də duy pə baə ki* ‘in their garden’). Such a different position of the preposition is natural in related languages, since in Proto-Indo-Aryan, Proto-Indo-Iranian, and even in the later stages the position of function words was not fixed, and the fixation of their position occurred relatively late.

However, in individual languages of the first type – in Yidgha and the majority of languages of the Shughni-Rushani group – we see cases in which the preposition is placed inside the possessive construction (with its gradual transformation into a prefix). Such cases are only a few, and they are all associated with inalienable possession: with the names of body parts, with sides (front, on the side, etc.), and with certain concepts that are inalienable from the possessor (memory, mind, sleep, etc.). Cf.: Sh. *mu tar dūst* ‘in my hand’; *xu tar bīst* ‘in self’s pocket’; *mu ba yođ* ‘in my memory’; *tu pi pođ* ‘on your leg’; *wi pi tanā* ‘on his body’; Ru. *xu par cāmēn* ‘with one’s own eyes’; Bt. *xu tar dōst* ‘in his hand’; Yidgha *mən tra zil* ‘in my heart’ (cf. Shughni *tar xi cīd* ‘in one’s house’; Yidgha *də x(w)oi kyēi* ‘in self’s house’). This construction for the modern languages is an archaism and is often violated (although we find cases of the lexicalization of the fusion of nouns with such ‘prefixes’ and also their transformation into postpositions. Cf.: Ru. *ǰār pa-kāl* ‘above the city’; Yd. *nā-mən da-pīr* ‘in front of me’). Such constructions are found in similar languages of the area.

The development of the opposition between alienable/inalienable possession through different positions of the preposition can be explained by the fact that in former times – in the Proto-Shughni language and in earlier stages of a series of other languages – terms of inalienable possession were used invariably with enclitic pronouns, and for this reason a prepositional modifier expressed as a full pronoun (or noun), became proleptic and was not included into the syntactic group in question and was perhaps treated as an oblique object: ‘At-me in my hand’. Precisely such constructions with terms of inalienable possession are found in other languages of the area (for instances, in Shumashti). The remaining nouns appeared either as enclitics or as prepositional modifiers – full pronouns or nouns, which here were already not proleptic, but were

---

<sup>20</sup> Maspero 1948:175; Andronov: 250-256; Zograf: 156

rather components of a modified group which functioned as a true syntactic group . . . Such a position was preserved in the majority of Shughni-Rushani languages (except for Sarikoli, where the second type became generalized) and in Yidgha even the disappearance of adnominal enclitics (in Waigali the preserved adnominal enclitics adjoin to the nominal stem, anticipating the case formant, cf. the reverse order in Munji, where the first type of construction was generalized).

The obligatoriness of the enclitics in terms with inalienable possession and their cohesion with the stem – i.e. the expression via enclitics of the corresponding grammatical category – does not fall out of the proto-language system, but rather has a direct analogy in Burushaski. Here, the clear opposition between alienable/inalienable possession is transmitted via the optionality/obligatoriness of the proclitic markers of person and (in the case of the third-person) the class of the possessor (cf. the proclitics in the Prasuni language). Here, constructions with a proleptic full pronoun of the type *ja a-rin-ulo* (lit. my my-hand-in) are typical. Analogous constructions in a substrate language could have facilitated the establishment of constructions of this type in these languages.

It is interesting that this same presence of the opposition of different types of possession leads us back to the opposition of animacy/inanimacy, as it can only be inherent for animate nouns (especially if we consider that terms of relation, which are often attached to animals as well, are formed here with the first type – that is, as a concept of alienable possession).

In this linguistic area, we find a tendency which is unusual for Indo-Iranian languages in the reorientation of the category of gender from a purely grammatical base to a semantic base. Usually, the Proto-Indo-Aryan category of gender in modern Indo-Iranian languages either disappears, or is preserved in the form of the same grammatical base, which is based on particular nouns, as it was in the ancient period, as well as the vocalization or the end of a word (albeit with certain changes in different languages). However, in the modern North Pamir group we see obvious deviations: in Yazghulami this category was formed entirely upon a semantic base and became essentially a semantic class: all nouns which refer to men and to objects, without regard to the type of stem, belong to the so-called ‘masculine’ gender, whereas nouns which refer to women and all animals, regardless of sex, belong to the ‘feminine’ gender (the term ‘gender’ here has a purely etymological sense – ‘masculine’ was formed through the merger of the historical masculine case with the neuter (inanimate), and the feminine absorbed the names of animals, including those such as *yew* ‘bull’, *bəč* ‘he-goat’; etc.). The category of gender has evolved even in the neighboring Shughni-Rushani group. With the preservation here of the older concept of division into gender classes based on the vocalization of noun stems, we also find deviations related to the intersection of the category of gender for inanimate nouns with that of individual/general, concrete/abstract; etc. Nouns which indicate an inanimate subject in these languages and which belong to the feminine gender (based upon their type of stem), change to masculine gender whenever they appear as the generalized name for the group of homogenous objects: cf. Ru. *dum māwn mu-r dāk* ‘give me that apple (fem.)’ but *day māwn tar bozor yōs* ‘take those apples (lit. that apple – mass) to the marker’.<sup>21</sup> There are also other semantic criteria

---

<sup>21</sup> Sokolova 1959: 108; 1966: 371. Karamshoev 1979: 28.

which are at play in ascribing nouns to one gender or another.<sup>22</sup> Thus, a series of nouns, in ambiguous conditions regarding their semantics, can change their grammatical gender without changing their form. For instance, while there is a general tendency for the names of cultivated plots of land to belong to the feminine gender, in Shughni and Bartangi the nouns *māḡ* ‘peas’; *žindam* ‘wheat’, Sh. *čūšč*, Bt. *čöšč* ‘barley’, when belonging to the general meaning of grain, are masculine, but when indicated that which was sown in a given field, belong to the feminine gender.<sup>23</sup> There is a series of other examples of nouns which belong to both genders.<sup>24</sup>

Essentially, this process signifies the reorganization of the gender distribution of nouns from a formal basis to a semantic (i.e. class) basis.

The division of nouns not by gender, or by the outer appearance of a stem – or, at least, not only on this basis – was typical of the pre-Indo-European languages of the region. This kind of system can be seen in Dravidian languages<sup>25</sup> and in Burushaski. Nonetheless, in modern Dravidian languages the basis for the classification system for nouns is the semantic criteria of person/non-person, with a further subdivision of nouns with a human meaning into masculine and feminine genders). The same principle for the division of nouns, expressed in the type of agreement which a noun receives and in corresponding interrogative pronouns, is found in practically all living Iranian languages.

In Burushaski, the four-class system – I masculine; II feminine; III animals and individual objects; and IV – remaining objects, material, and abstract concepts – results in a more complex scheme of oppositions. Here we observe, in particular, cases in which nouns belong to two genders, and the selection of a particular gender (either III or IV) in individual cases depends on semantic nuances (individual vs. general, object vs. material or abstract concept; fruit vs. tree; etc.). For instance: *yašil* ‘stick; cane’ is III, but the same word meaning ‘wood/firewood’ is IV; the word *bālt* ‘apple’ is III, but the same word meaning apple tree is IV; *baiyu* ‘rock salt’ is III, but the same word meaning ‘grain salt’ is IV; *bayundo* ‘bread from fermented dough’ is III, but the same word meaning the dough itself is IV. This kind of system resonates with the with the analogous phenomenon with inanimate nouns in the Shughni-Rushani group. In other words, the classification of a series of inanimate nouns to the feminine (or masculine) in the Shughni-Rushani group is, as with the classification of analogous nouns to Class III (or IV) in Burushaski, is a sort of semantic marker (meaning concrete vs. abstract, object vs. abstract concept, individual vs. general, etc.) and is used to indicate the class to which a particular noun belongs in a particular context.

In Burushaski we also find the opposition between (human/non-human?), which is reflected in the first two classes, and already in a tenuous way we can see the old opposition of animate/inanimate, which is observed in the following way: Class III now includes, in addition to the names of animals, also the names of fruits, trees and their parts, parts of the body and names

---

<sup>22</sup> See the lists of words divided into gender in concrete descriptive works by D. Karamshoev, N. Karamkhudoev, Kh. Kurbanov, M. Fayzov, etc. For a summary see Karamshoev 1979: 99-100.

<sup>23</sup> Karamshoev 1963. *Bajuwi dialect of the Shughni language.*; Karamkhudoev: 59

<sup>24</sup> See Karamshoev, *Kategoriya roda*: 27.

<sup>25</sup> On the connection between gender and in individualness/generality, the opposition in Proto-Dravidian between animacy/inanimacy, and other semantic groups, see N. V. Gurov



of some celestial and natural phenomena – thus, a word which, although it is not a human, belongs to the class of living things, for instance a living cycle. At the present time, we can include here a group of other words, though, it is believed that they came about later, with a primary meaning belonging to Class III – namely the class of living beings, plants, and natural phenomena – interpreted primarily as living things, in opposition to Class IV, which combines names of objects and concepts of the non-living realm.<sup>26</sup> Thus, the fundamental opposition in Burushaski is the opposition between non-living and living things.

It is known that the opposition of human/non-human in the modern Iranian languages and a part of the Indo-Aryan languages, built on the feature of social activity, is a relatively late development: it is preceded by the opposition of animacy/inanimacy, which is attested in Proto-Indo-Aryan through differing nominal paradigms. And the fact that the reconstruction of the gender system in North Pamir languages has emerged (in Yazghulami, first and foremost, but also in the Shughni-Rushani group as one of several processes) through the isolation of names of animals from names of inanimate objects, once again attests to the presence here of an implicit distinction between animate/inanimate.

It is natural that not all purportedly substratal features of the languages of this area have been identified – as concerns semantic categories the task is particularly difficult, considering the fact that most of the languages in question are not well studied and material on them is difficult to access. Undoubtedly, in due course we will have such features, as well as a list of the languages which have them. Nonetheless, even the short overview of ‘anomalies’ given here allows us to make some conclusions. The features discussed here – neither the formal features nor the content features – are not found simultaneously all together in any of the Indo-Iranian languages of the region and consequently do not form a real system at the present time. They also do not fall out necessarily from the patterns of the spontaneous development of the Proto-Indo-Aryan structure and, consequently, may be the result of changes which were brought about by corresponding structures of the linguistic substrate.

The totality of formal structural changes brought about by the substrate gives us somewhat of a glimpse into the structure of the specific subsystems of the substrate (its phonological make-up, the structure of its numerals, pronouns, the word-forming models of its nominal system, and lexical elements which are constrained to the reality of the speakers, etc.).

The totality of the anomalies found in the substantive character (of modern languages) indicate that they are all linked to a rather consistent, logical system based on the factor of active/inactive with some variations: regarding the general living/non-living or animate/inanimate and ‘active/inactive in a given situation’. If we assume that the system in question was inherent to the linguistic substrate, then we must acknowledge that the pre-Indo-European languages which existed in this region are reminiscent of active-stative languages, in which the feature of active/inactiveness is projected at various levels.

Indeed, traces of the division of verbs not based upon transitivity/intransitivity, but rather on animacy/inanimacy of the actor and differing lexemes or copular forms for animate and inanimate subjects speaks to the opposition of the verb in this substratal system precisely along

---

<sup>26</sup> Lorimer. *The Burushaski Language*: 20-25.

these lines. The presence of verbs indicating a subject's quality (of the type 'be tall') are very common for languages of an active-stative nature, in which the independent lexical category of 'adjective' is lacking. The opposition between alienable/inalienable possession has a clear expression in languages of the active-stative type, where it distinguishes animate nouns (which have inalienable possession) from inanimate nouns (which do not). The observation that in a number of languages in the region there are differing paradigms for nouns with meanings of nouns with animate referents, on the one hand, and inanimate referents, on the other (not treated here due to a lack of space), can be explained through the hindering influence of the substrate, which led to the preservation in some period of the opposition which took place in the Indo-European period and which continued into the Proto-Indo-Aryan period.

The presence of noun classes is also typical for languages of the active-stative type, where they are one of the most substantial implications for activeness, dividing nouns into either active or non-active classes (with certain further modifications, such as those which were discussed above). In this regard, Burushaski, with its class system and primary opposition between living/non-living, is closer to the prototype than Dravidian. And finally, the opposition between inclusive/exclusive or inclusive/non-inclusive is characteristic for this type of modification, where the narrator indicates the situational activity/inactivity (or inclusivity/exclusivity) of the talking partner.

Therefore, the impression emerges that the substrate must have had features of active-stative typology. In this regard, it can be matched most closely to Burushaski<sup>27</sup> and Himalayish languages, which still today display features of active-stative structure or its formal elements. In some cases, similarities in these types of characteristics are seen in modern-day nominative-accusative Dravidian languages, though the effect they may have had on Indo-Iranian languages in the presently discussed group could have been lesser (or mediated). If we consider that the Proto-Indo-Aryan (or Indo-Iranian?) period, judging by what we see in Old Iranian and Old Aryan languages, were characterized by a nominative-accusative typology, in which animacy/inanimacy was manifested residually, in the periphery of the nominal paradigm, then the development of central and peripheral features of active-stative typology in the languages of the region during the subsequent periods could have hardly been spontaneous.

The relatively late emergence or display of substratal features is not something out of the ordinary which is characteristic only of this region. It is noted in many other regions of the world and is characteristic for those cases in which the modern-day languages of the region develop those tendencies which long before could have been catalyzed or even put into place by the substrate.<sup>28</sup> It is possible that substantive elements, being very sturdy, can be preserved for quite a long time in the unwritten realm and break through into the written language relatively late.

From all that is written we can make one more conclusion: the typological anomalies in the common structure of the genetically related languages of the area offer valuable material for the

---

<sup>27</sup> On the substratal role of Burushaski see also: LSI, VIII, pt. 2. Calcutta 1919: p. 6. Also: Abaev 1949:589.

<sup>28</sup> See: V.I. Abaev. On the linguistic substrate: 63-64; The typology of Armenian and Ossetian: 27; V. N. Yartseva: 20-21; etc.

establishment of the structure of the substrate languages not only with respect to particular subsystems, but also regarding holistic typological characteristics.