

On some Romanian substrate words with Albanian comparanda: etymologies and implications

The paper offers a new angle on the etymologies of some Romanian substrate words with close Albanian comparanda, focusing on distinguishing each word's language of origin. Two new Albanian loanwords are identified, *berc* 'with a short tail; tailless' and *driză* 'hairy crab-grass', while *rangă* 'crowbar' and *urdă* 'a kind of cheese similar to ricotta' are identified as Thracian borrowings. Further, *groapă* 'pit' (< **g^hrob^h-eh₂*) is attributed to a minor substrate that preserved PIE **o* — Paeonian or Brygian. Finally, *mărar* 'dill' is derived ultimately from Ancient Greek **maráthrion*, a diminutive of *márathron* 'fennel', through the mediation of both Thracian and Albanian.

Keywords: Romanian language; Romanian etymology; Paleo-Balkan languages; Albanian language; Thracian language.

1. Introduction

Romanian preserves a considerable number of words from Paleo-Balkan languages — the poorly attested languages native to pre-Roman Southeastern Europe. These words were acquired largely through the shift of Paleo-Balkan speakers to Latin and account collectively for the vast majority of the Romanian lexical substrate.¹ Currently, there is no consensus on the linguistic stratification² or periodization of this vocabulary,³ and the problem of the Romanian substrate is intimately connected to the yet-unsolved problem of the Romanian Urheimat. Based on historical considerations, substrate words may have entered Balkan Latin as early as the 2nd century BC, when Latin began to spread in the interior of the Balkan Peninsula following the Third Macedonian War (171–168 BC), and as late as the final quarter of 1st millennium AD, prior to the cessation of Romanian-Albanian⁴ contacts and the extinction of Daco-Moesian and Thracian.⁵ This lengthy period of often intensive contacts with Paleo-Balkan languages⁶ accounts for the significant influence that these languages exerted upon Romanian.

¹ Greek, usually not classed as a Paleo-Balkan language due to its extensive attestation (see Нерознак 1978: 7), also contributed to the Romanian substrate: Romanian exhibits a layer of early Greek loans of substrate typology — particularly names of small animals, such as *cărăbuș* 'cockchafer', *drepnea* 'blackswift', and *mormoloc* 'tadpole' (Diculescu 1927: 486–489). Additionally, any influences acquired from languages of non-Balkan origin that predated Latin in Southeastern Europe (e.g., Sarmatian) may be regarded as substratal.

² Research aimed at identifying layers within the Romanian lexical substrate has been limited. An early attempt, now outdated, was Philippide (1927: 695–743). Later, Ivănescu (1980: 251–257) distinguished an Albanian and a non-Albanian layer in the Romanian substrate, but without providing an extensive analysis.

³ According to Ivănescu (1980: 171), the bulk of Romanian substrate words entered the language between 450 and 700 AD. Other authors propose earlier dates: Paliga (2024: 43–44) asserts that Thracian influenced Balkan Latin until the 6th century AD.

⁴ Albanian, regarded by Matzinger (2018: 1790) as the continuer of an undocumented Balkan idiom distinct from both Illyrian and Thracian, is the sole modern descendant of a Paleo-Balkan language.

⁵ Daco-Moesian and Thracian appear to have been still in widespread use at the time when the Slavs arrived in Southeastern Europe. See Yanakieva (2018: 59–61) on the situation of East Balkan languages in Late Antiquity.

Among the Romanian substrate words, particular attention has been given to those with close parallels in Albanian.⁷ However, despite more than 160 years of research efforts, the problem of identifying and etymologizing Romanian substrate words with Albanian comparanda remains unsolved.⁸ The identification of these words is complicated by the very late attestations of both languages⁹ and the heterogeneity of the shared Romanian-Albanian vocabulary.¹⁰ Among Romanian substrate words common with Albanian, reliable etymologies are largely limited to those analyzable as native in Albanian, of which some evidence Albanian-specific developments,¹¹ whereas others do not.¹² Several important Romanian substrate words shared with Albanian remain with no plausible Albanian etymology.¹³

In recent decades, it has often been assumed that this group of Romanian substrate words largely derives from a single language (Nagy 2003: 194). This presumptive source language has been regarded as ancestral or otherwise very closely related to Albanian (Brâncuș 2022: 46; Matzinger 2018: 1790; Paliga 2024: 46; Poghirc 1969: 314; Slușanschi 1981: 340). However, the difficulty in providing Albanian etymologies for a significant share of these words suggests that they did not originate in a single language or an “Albanoid” group¹⁴ of closely related languages.

The linguistic diversity of the ancient Balkans offers a compelling alternative to the Albanoid hypothesis as an explanation for the Paleo-Balkan vocabulary shared by Romanian and Albanian. Multiple languages were spoken in pre-Roman times across the territory between present-day Romania and Albania, including Daco-Moesian, Thracian, Paeonian and Brygian.¹⁵ While Albanian has itself been often suggested to continue either Daco-Moesian or Thracian (Brâncuș 2013: 419; Georgiev 1981: 132–140; Paliga 2024: 46; Witczak 1995: 310–311), it markedly differs from these languages in its treatment of several PIE consonants and clusters, particularly *sC groups (Kaluzhskaya 1993: 70–71; Ködderitzsch 2001: 82), rendering such affiliations implausible. Albanian likewise diverges from Paeonian and Brygian, having shifted PIE *o to *a in pre-Roman times (Matzinger 2017: 39–40), whereas the latter languages both preserved it (Duridanov 1985: 134–135; 1993: 64–65).

Given these considerations, the essential principle guiding our investigation is taking into account the diversity of possible sources of the shared Paleo-Balkan vocabulary of Romanian and Albanian. This paper follows this framework to analyze six such words, offering arguments for their attribution to a particular language rather than to a generic substrate.

⁶ See Rusakov (2017: 556–557) for information on the intensive Romanian-Albanian contacts, discernible at multiple linguistic levels.

⁷ An early discussion of this subject is available in Miklosich (1861: 6–10).

⁸ Reflecting this uncertainty, Brâncuș (2005: 48–50) estimates that Romanian has between 90 and 150 substrate words shared with Albanian. New candidates continue to be proposed, such as Rom. *îngheșui* ‘to cram, jam’, linked by Schütz (2002: 19) to Alb. *gjesh* ‘to knead’.

⁹ Albanian and Romanian are virtually undocumented for most of their history. Substantive attestations first emerge in the 15th century AD for Albanian (Rusakov 2018: 1716) and around 1500 for Romanian (Maiden et al. 2020: 3–4).

¹⁰ Romanian and Albanian also share extensive Latin and Slavic vocabulary layers besides the Paleo-Balkan one.

¹¹ One such example is Rom. *vatră* ‘hearth’ < Alb. *vatër* id., whose /v/ arose via the Alb. *ā- > *w- sound change in the earlier *ātrā ‘fire’ (Schumacher 2009: 46).

¹² An example is Rom. *gard* ‘fence’ ~ Alb. *gardh* id. < *g^hord^h-o- < PIE *g^herd^h- ‘to enclose, gird’ (Schumacher 2013: 211). The word formation is straightforward and the phonetic developments are not diagnostic: PIE *o > a and *g^h, *d^h > g, d have also been attributed to Dacian and Thracian by Duridanov (1995: 45–46).

¹³ An example is Rom. *copil* ‘child’ ~ Alb. *kopil* ‘bastard’. The main attempt at an Albanian etymology starts from Alb. *pjell* ‘to give birth (of animals)’ and assumes a *ko-* prefix and an ‘of bad birth’ original meaning (Hubschmid 1983: 507–510). The speculative word formation and phonetic difficulties render this etymology unpersuasive.

¹⁴ See a brief presentation of the Albanoid hypothesis in Friedman and Joseph (2025: 33–34).

¹⁵ See Duridanov (1995: 45; 2000: 23–25) and Hill (2011: 411).

2. Etymologies

2.1. Rom. *berc* ‘with a short tail; tailless’

This word has often been considered of unknown etymology (DLR 1913–1949, 1, pt. 1: 545; Șăineanu 1929: 62; Tiktin 1903–1925, 1: 177).

The explanation from a surmised Lat. **brevicus* < *brevis* ‘short’ (Bogrea 1921: 251) requires an arbitrary metathesis. Moreover, other Romance languages do not validate **brevicus*, and the adjective-forming Lat. *-icus* suffix has not been historically productive in Romanian.

The derivation from OHG *brecha* ‘breaking’ (Scriban 1939: 164) is semantically vague.¹⁶ Formally and distributionally, the hypothesis presents further difficulties, hinging upon the same unexplained metathesis and requiring an unlikely OHG *x* > Rom. *k* outcome¹⁷ as well as Old High German-Romanian contacts, which are not evidenced. Hypothetical mediation through an unattested Slavic form (Ciorănescu 2001: 88) fails to resolve these difficulties.

Substratal etymologies have also been proposed. Vinereanu (2023, 1: 240) derives it from PIE **mreǵh-* ‘short’ through Dacian, with an unexpected metathesis and velar stop.¹⁸ Paliga (2024: 187) regards *berc* as a probable substrate word from a surmised *bar-/ber-* root extracted from toponyms and presumed to mean ‘hill’. However, such a root is not validated, and the semantic link to ‘hill’ is not compelling.

Saramandu (1987: 126–130) connects *berc* with Rom. *bercă* ‘sheep with short and curly wool’¹⁹ and Alb. *berr* ‘sheep; goat’, but regards it as a diminutive of a lost **ber* through an earlier **berce*, with an unclear suffix. Saramandu (1987: 130) further suggests a ‘sheep with short and curly wool’ > ‘with a short tail; tailless’ semantic evolution, which is unconvincing.

We propose instead that *berc* was borrowed from Alb. *berk* ‘goat’. The Albanian word is an internal diminutive of Alb. *berr* id. (Orel 1998: 21).²⁰ Early Romanians may have encountered *berk* as a term used to describe goats that resembled sheep in some way. A conspicuous difference between goats and sheep is that goats have much shorter tails than sheep; the historical importance of pastoralism among Romanians favored the generalization of an initial meaning ‘sheep with a short tail’ to ‘animal with a short tail’.²¹ Another conspicuous difference between goats and sheep is that goats have much shorter wool than sheep; this explains the meaning of *bercă*.²² For the semantic shift from the name of an animal to a perceived attribute of it, cf. Hun. *ravasz* ‘sly’ < ‘fox’ (Kálmán 1972: 56).

Finally, it is worth noting that *berk* is a word of low distribution in Albanian — Orel (1998: 21) mentions it is used only in *era berk* ‘stench of a goat’. However, given the large temporal gap between the age of intensive Romanian-Albanian contacts, dated to the mid-1st millen-

¹⁶ One could, at most, surmise a connection with tail docking, a practice performed in farm animals such as pigs, sheep and cattle (Sutherland and Tucker 2011: 180).

¹⁷ Such an outcome appears sporadically in auslaut in Old Church Slavonic loans, but the more frequent outcomes of OCS *-x* are Rom. *-f*, *-v*, *-u* (with dialectal variation), and *-h* in learned borrowings (Petrovici 1948: 190–192).

¹⁸ PIE **ǵh* > Dacian *z* (Ködderitzsch 2001: 82).

¹⁹ Rom. *bercă* lent Hun. *birka* ‘sheep; lamb, mutton’, Slovak *birka* ‘short-wooled sheep’, dialectal (Moravian) Czech *burka* ‘sheep’ (Pușcariu et al. 1926: 279), dialectal Russian *birka* ‘lamb, sheepskin’. Zaicz (2006 81) instead considers Hun. *birka* of either native or Slavic origin, but without explaining its word formation.

²⁰ Related to PSl. **bôrvъ* ‘boar, ram, wether’ (Orel 1998: 21) < PIE **b^hor-wo-* (Derksen 2008: 58).

²¹ Another semantic shift attributed to the historical predominance of pastoralism among Romanians is Rom. *făt* ‘boy’ < ‘young animal’ (Candrea and Densusianu 1907: 91–92).

²² The feminine gender of Rom. *bercă*, in contrast with the masculine gender of Alb. *berk*, may reflect the influence of the feminine gender of Rom. *oaie* ‘sheep’.

nium AD (Rusakov 2017: 557), and the first consistent attestations of Albanian, this circumstance does not strongly challenge our etymology. It may be that *berk* enjoyed wider usage in the mid-1st millennium AD, whether in Albanian as a whole or in a variety thereof that experienced greater contact with the Common Romanian dialect ancestral to Daco-Romanian. The absence of *berc* in Aromanian and Megleno-Romanian supports the latter scenario, suggesting that the word was borrowed from a northern Albanian variety which may have since disappeared through Romanization or Slavicization.

2.2. Rom. *driză* ‘hairy crabgrass’

MDA (2001–2003, 2: 226) provides no etymology for this rare word. DELR suggests an untenable derivation from German *Waldhirse* ‘foxtail millet’ through **valtirză* and **valdirză*.

The word is a transparent loan from Alb. *drizë* ‘thornbush’. The Albanian word derives from **dri-* ‘bush’, a root shared with AGr. *dríos* id., with the Albanian diminutive suffix *-zë*.²³ See the discussion of *drizë* in Demiraj (1997: 146).

2.3. Rom. *groapă* ‘pit’

There is little doubt that this word continues an o-grade result noun from PIE **g^hreb^h-* ‘to dig’ (LIV 2001: 201–202). The comparison to Alb. *gropë* id., German *Graben* ‘ditch’, OCS *grobŭ* ‘pit, hole’ was already made by Miklosich (1861: 9). What remains is to identify the language of origin.

A Slavic (Cihac 1879: 129–130; Georgiev 1965: 76; Ungureanu 2016: 210), Thracian (Paliga 2024: 307)²⁴ or Dacian (Vinereanu 2023, 1: 704) origin cannot explain the /p/, while a borrowing from OHG *crôpa* ‘pit, cavity’ (Orel 1998: 124)²⁵ is distributionally impossible.

DLR (1913–1949, 2, pt. 1: 315) derives *groapă* from Albanian, which does not suffice to clarify the history of the word. Regarding the /p/, Hyllested (2016: 73–77) brings compelling arguments for a syllable-final **-b- > *-p* development in early Albanian, occurring after the apocope of the unstressed vowel. Since the **-p* in the new word-final position was retained after vocalic suffixation, such as with the feminine suffix *-ë* (Hyllested 2016: 74), Albanian can account for the **b^h > p* outcome. However, *gropë* remains difficult to explain as a native Albanian word. This is because Alb. /o/ does not continue PIE **o*, but PIE **eh_{1/2}* (Demiraj 1997: 41–45), which **g^hreb^h-* cannot provide. If starting instead from **g^hrēb^h-*, /o/ would be expected, but this requires an arbitrary long grade.

A Slavic > Albanian > Romanian borrowing scenario has been suggested by Meyer (1891: 131) and Tiktin (1903–1925, 2: 703). Matzinger (2006: 213) posits that the early date of the borrowing explains the surmised *b > p* outcome. Yet this hypothesis is seriously marred by

²³ On this suffix, see Arapi (2012: 134). Other Romanian words borrowed from Albanian suffixed with *-ză* include *coacăză* ‘currant’ (Çabej 1964: 73) and *pupăză* ‘hoopoe’ (Orel 1998: 350).

²⁴ Paliga claims that PIE **b^h* yielded Thr. *p*, giving Rom. *jupân* ‘master’ and *stăpân* ‘master, owner’ as examples. However, *jupân* is not of Thracian but of eastern origin, related to Bulgar *zoapan* (Blažek 2006: 515) and Old Uyghur *čopan* ‘chancellor’ (Ramstedt 1951: 72), borrowed into Romanian via Proto-Slavic (Leschber 2021: 213). Cf. further related official titles: Chinese *chù bàn* (处半), Bactrian *sōpano*, Khotanese *chaupaṃ* (Zhan 2014: 77). Regarding *stăpân*, Philippide (1927: 14–15) compares it to ARo. *stăpuescu* ‘I stop, I master myself’ and further to Sanskrit *sthāpayati* ‘to make to stand’, implying a derivation from PIE **steh₂-* ‘to stand’ and an original meaning of **‘one who makes (things, people) stay put’*. The **-p-* suffix is unclear, but a hypothetical earlier **-b^h-* form is unevicenced. See a discussion of the origin of the possibly related Indic *-pāya-* causative suffix in Khoshsirāt and Byrd (2023).

²⁵ A variant of OHG *gruoba* ‘pit, cavity’ (Schade 1872–1882: 355) < PGm. **grōbō* ‘hole, pit’, an ō-stem derived from PGm. **graban* ‘to dig’ < PIE **g^hreb^h-* id. (Kroonen 2013: 191; 185–186).

chronological difficulties. No loanwords in Albanian showing the $*-b- > *-p$ outcome have been identified so far: the /b/ is preserved in Slavic loans *kob* ‘treachery’, *rob* ‘serf’, *cub*¹ ‘short-tailed; with the tail cut off’, *udob* ‘easy, convenient’ and even Gothic loan *cub*² ‘robber’ (Hyllested 2016: 67). If *gropë* is indeed of Slavic origin, the borrowing must have taken place very early. One may surmise that early Albanians borrowed *gropë* from Proto-Slavic just before the Alb. $*-b- > *-p$ shift completed, and shortly thereafter borrowed *cub* from remnant Germanic populations of the Balkans. However, the earlier form of PSl. $*grǫbb$ ‘grave’ was $*grabu$ (Klotz 2017: 112), and the PSl. $*a, *u > *o$, $*b$ shifts only occurred after 600 AD (Leschber 2021: 208). Unless the borrowing of *cub*² from Gothic was as late as the 7th century AD, relative chronology demands that we start from $*grabu$. This, in turn, would require an unevidenced early PSl. $*a > Alb. o$ outcome.

These difficulties suggest the need to move beyond the traditional hypotheses. The final-syllable $*-b- > *-p$ shift has a notable Balkan parallel. The name of *Astibus*,²⁶ a town of Paeonia, an ancient region centered on the middle reaches of the Vardar (Axios) River (Duridanov 1985: 133), has been preserved in Macedonian as *Štip*. Displaying the loss of anlaut unstressed *a* characteristic of Albanian, the modern name has been regarded as transmitted by Albanian (Barić 1955: 49). While its $b > p$ shift may have occurred in Albanian, an alternative possibility is that it occurred already in Paeonian²⁷ and the toponym was passed to Albanian as $*Astip$. This circumstance warrants consideration of the possibility that *groapă* originated in Paeonian.

Among the distinguishing features of Paeonian proposed by Duridanov (1985: 134–135) we remark the one needed for explaining *groapă*: the preservation of PIE $*o$, based on the Paeonian hydronym *Pontos* (< PIE $*pent-$ ‘find a path’ through the o-grade, related to Gr. *póntos* ‘sea’ (Янакиева 2009: 112)) and ethnonym *Stoboi*. The preservation of PIE $*o$ is not exclusive to Paeonian among the eastern Paleo-Balkan languages, however: it also characterizes Brygian, another variety historically spoken near the Axios River (Duridanov 1993: 64–65).

Therefore, Rom. *groapă* and Alb. *gropë* continue a Paeonian or Brygian reflex of $*g^hrob^h-eh_2$. Most likely, Albanian mediated the Paeonian or Brygian word to Eastern Romance, as suggested by the Albanian transmission of *Astibus*.

2.4 Rom. *mărar* ‘dill’

Long linked with Gr. *mámathron* ‘fennel’ and Alb. *maraj/mërajë* id. (Capidan 1922: 458), this word is usually regarded as substratal, with no etymological solution provided (Boerescu 2018: 52; Brâncuș 1983: 96–97; Poghirc 1969: 345). The Greek word was attested in Mycenaean as *ma-ra-tu-wo*, and may have been borrowed from Pre-Greek (Beekes 2010, 2: 903–904).

Diculescu (1927: 460) proposes a borrowing from Gr. $*maráthrion$, a reconstructed diminutive of *mámathron*. However, the Gr. *th* > Rom. $\emptyset$ outcome would be unexplained.

Orel (1998: 244–245) regards *mărar* as borrowed from Alb. $*mararja$, an earlier form of *maraj*. This is a plausible viewpoint; for Alb. *rj* > *j*, see Schumacher (2013: 247). The obsolete Rom. *mărariu* ‘dill’ (Boerescu 2018: 52) and ARo. *măral’u* ‘dill’ (Capidan 1922: 458) buttress this hypothesis: the Alb. *rj* > *j* sound change proceeded with a λ intermediary that survives in some Tosk varieties (Schumacher 2013: 247). Thus, *mărar* and *măral’u* were borrowed from Albanian, with the former an earlier loan than the latter.

Next, we will elucidate the link between *maraj* and *mámathron*. First, the reconstructed diminutive $*maráthrion$ is necessary to account for the $*rj$ in the history of the Albanian word.

²⁶ Derived by Georgiev (1981: 172) from $*ad-tib^hos$ ‘at the marsh’.

²⁷ Paeonia was classed as a distinct ethnic entity from Thrace, Illyria and Macedonia by some ancient writers, though conflicting reports describe Paeonians variously as Thracians, Phrygians or Illyrians (Katičić 1976: 117).

The loss of **th* can be explained through a **drj > *rj > j* development, well-known in Albanian,²⁸ entailing an earlier **maradrjV-*, and an Alb. **d ~ Gr. th* correspondence. To account for it, Huld (1986: 247–249) proposes that *maraj* was borrowed from a Macedonian cognate of **maráthrion*,²⁹ putting forward two other loanwords in Albanian with a putative Macedonian treatment of PIE voiced stops:

- Alb. *mëllagë* ‘mallow’ ~ AGr. *malákhē* id.
- Alb. *bredh* ‘fir tree’ ~ AGr. *bráthu* ‘savin’.

But these examples fail to evidence an original voiced aspirate. AGr. *malákhē* can be compared with Semitic words like Hebrew *malúakh* ‘saltbush’,³⁰ and the Albanian and Ancient Greek words may be independent non-Indo-European loans. Regarding AGr. *bráthu*, comparanda such as Aramaic *bərāt* ‘cypress’ and Assyrian *burāšu* id.³¹ do not indicate an original **d^h*. Furthermore, Alb. *bredh* is rather a native word from PIE **b^herHǵ-o-* ‘birch’ (Demiraj 1997: 107–108), and therefore it must be separated from *bráthu*. Thus, another explanation for the Gr. *th ~ Alb. *d* correspondence is needed.

It may be more appropriate to start from **maráthrion* rather than from a postulated cognate of it. As fennel and dill are aromatic herbs — a category particularly susceptible to lexical borrowing through trade relations — the transmission of Gr. **maráthrion* to Alb. **maradrjV-* may have occurred through an intermediary language.

A plausible intermediary language can be identified. In Thracian, voiceless stops were voiced dialectally when adjacent to **r* or **n* (Duridanov 1985: 117–118; Georgiev 1983: 1171–1172). Remarkably, evidence from Romanian confirms this development. In Rom. *Medru* ‘Demetrius’ and *Sumedru* ‘Saint Demetrius’, the *tr > dr* shift, thought unexplained (Ivănescu 1980: 242) can be attributed to Thracian mediation. Regarding the specifics of these dialectal Thracian developments, Duridanov (1985: 118) proposes Thr. **t^hn > *dn* and **nk^h > *ng* shifts. However, it seems more typologically plausible that voiceless aspirates conditionally evolved to voiced aspirates, subsequently represented as plain voiced stops in the Greek and Latin alphabets as these lacked dedicated symbols to render these sounds. While this issue bears no direct relevance to Romanian etymology as Romanian does not distinguish voiced aspirates, we will opt to represent the Thracian sounds in question here as voiced aspirates.

We propose, thus, that a Thr. **marad^hrjV-* was borrowed from Gr. **maráthrion* and subsequently reached Albanian, where it yielded **maradrjV-*. A later version of **maradrjV-* subsequently lent Eastern Romance *mărar* and *măral’u* before evolving to *maraj/mërajë*.

2.5 Rom. *rangă* ‘crowbar’

This word has generally been considered unexplained (see MDA (2001–2003, 4: 124)). Vinereanu (2023, 2: 365) tentatively suggests a formally untenable derivation from PIE **reik-* (‘to stretch up, to wobble; pole, bar’ (Pokorny 1959: 862)).

Semantically, a derivation from PIE **wronk-eh₂* (also reflected in PSl. **røkà* ‘hand, arm’ (Greenberg 2017: 529) and Late Lat. *branca* ‘paw’, a loanword (Pokorny 1959: 1155)) is quite satisfactory. The PIE noun derives from a verbal root **wrenk-* preserved in Lithuanian *riñkti* ‘to gather’ (Derksen 2008: 440).³²

²⁸ See Hamp (1976: 34).

²⁹ PIE **b^h, *d^h, *g^h > AGr. p^h, t^h, k^h, AMac. b, d, g* (Georgiev 1981: 167–169).

³⁰ Beekes (2010, 2: 896) reconstructs a **mal^wak-* pre-form.

³¹ See Beekes (2010, 1: 234–235).

³² Galician *ranco* ‘a tool used to clean ashes and embers’ appears to be another substrate word in a Romance language derived from **wrenk-*, paralleling *rangă* in its semantic evolution.

The Romanian word has a lookalike in Alb. *rangë* ‘housework’, plausibly derived from the same **wronk-eh₂* by Orel (1998: 365). Due to the wide semantic gap, Albanian cannot be the source of Rom. *rangă*, however.

Instead, the Thracian voicing of stops adjacent to **n* noted in the discussion of *mărar*, together with PIE **o* > Thr. **a*,³³ suggests Thracian provenance. While the attested Thracian material provides no reliable examples of the treatment of PIE **wr*,³⁴ Romanian evidence supports **r* as the Thracian reflex of PIE **wr*. Vinereanu (2023, 2: 380–381) derives the otherwise unexplained Rom. *rătie* ‘twig; brushwood’ and *rătină* ‘small branch’ from PIE **wreh₂d-* (‘root’, but perhaps also ‘branch’, as suggested by the Greek and Latin data (de Vaan 2008: 512–513)). Based on the implied PIE **d* > *t* outcome, attributable to Thracian (Duridanov 1985: 137; Georgiev 1981: 128),³⁵ *rătie* and *rătină* can be plausibly identified as Thracian loans.

Finally, we add that the preserved /a/ in *rangă* shows that the borrowing from Thracian occurred no earlier than the 6th century AD. This is because the Rom. *an* > *in* shift, characteristic of native words and very early loanwords, also surfaces in Rom. *jupân* ‘master’, an early Slavic loan. The fact that Thracian still influenced Romanian at this relatively late date can be explained through Thracian’s resistance to Latinization and Hellenization. As inferred from toponymic evidence, only a limited part of the Thracian population had adopted Latin or Greek at the time of the Slavic migration to the Balkans (Yanakieva 2018: 60–61).

2.6. Rom. *urdă* ‘a kind of cheese similar to ricotta’

Exploiting the fact that this culinary product is prepared by boiling, derivations from PIE **werH-* ‘to be hot’ (LIV 2001: 689) have been proposed for *urdă* and Alb. *urdhë* id. (Orel 1998: 487–488; Russu 1981: 412). Semantically, this is plausible, but the formal aspect needs elucidation.

If assuming an Albanian origin, the suffix has to be identified with Alb. *-dhë*, but this is not comfortable, because the suffix only appears in a few words (see Ajeti (2017: 531–532)) and does not generally attach to verbal roots. A stronger objection to this scenario is that the *ur-* reflex cannot be explained through Albanian if starting from **werH-*. Orel (1998: 487) assumes either a zero-grade or an o-grade formation, but neither option can explain *urdhë*. PIE **C_rHT* yielded Albanian *CraT*, while the normal result of PIE **o* is Alb. *a*, with *u* appearing as an outcome after *k*, *g* and *h* (Schumacher 2013: 211, 213, 230).

These difficulties prompt us to explore derivations from **werH-* through a pre-Roman language other than Albanian. The *ur-* appears to reflect a zero-grade derivative. In this regard, Thracian offers an option, as its reflex of PIE **r* was **ur/or*, distinguishing it from Dacian, whose reflex was **ri* (Georgiev 1983: 1189). Moreover, laryngeals following a syllabic resonant

³³ See Georgiev (1983: 1165).

³⁴ Mihaylova and Mircheva (2022: 210) plausibly argue that Thr. *briza* ‘rye’, comparable to Sanskrit *vrīhi* ‘rice’, is a Wanderwort bearing little significance to questions of Thracian historical phonetics. It has been claimed, though, that Thr. *bria* ‘city’ is related to Tocharian A *ri* id. and Tocharian B *riye*, entailing a **wriyā* proto-form (Duridanov 1985: 107) and a PIE **wr* > Thr. *br/vr* shift. But the apparent Thracian-Tocharian isogloss is isolated, and the Tocharian words admit another etymology: Lubotsky and Starostin (2003: 264) propose a Chinese origin ultimately from Proto-Sino-Tibetan **riəH*, whence also Jingpo *məre* ‘town’.

³⁵ While the Thracian devoicing of the PIE mediae remains disputed (Yanakieva 2018: 47–48), some of the proposed examples have no alternative explanation. A well-known case is Thracian *sinupyla* ‘white briony’, contrasting with Dacian *kinouboila*³⁵ id. and Lithuanian *šūnobelė* ‘crabapple’, from a pre-form meaning ‘dog’s apple’ (Duridanov 1995: 48). Further examples have been advanced recently: Mihaylova and Mircheva (2022: 215–216) compare Thr. *torelle* ‘funeral song’ with AGr. *odūromai* ‘to wail, lament’.

elided before a consonant in Thracian, as shown from the onomastic element *-turma*, continuing **d_ṛh₂-mo-* < PIE **derh₂-* ‘to run’ (Mircheva and Mihaylova 2020: 57), while the PIE **w_r* > Thr. **r* shift discussed for *rangă* suggests an analogous PIE **w* > Thr. $\emptyset$ / $_$ **_ṛ* development.

The above evidence establishes **ur-* as the probable Thracian reflex of PIE **w_rH-*, making the language the likely source of *urdă* and *urdhë*. An additional factor speaking in favor of a Thracian origin is the suffix. Because of the Thracian voicing of stops near **r* also noted earlier, introducing a **-d^h-* suffix is unnecessary. Instead, the pre-form of the word can be reconstructed as a participle formation in **-teh₂*: **w_rH-teh₂*. Semantically, the formation is very close to Italian *ricotta*, etymologically ‘recooked’. Subsequently, the initial Thracian result of **w_rH-teh₂*, **urt^hā*, yielded **urd^hā*, which accounts for the discussed Romanian and Albanian words.

Regarding the pathway of borrowing, since Thracian is known to have significantly influenced Eastern Romance³⁶ but not Albanian, the Thracian word was most probably mediated to Albanian by Balkan Latin.

3. Summary and conclusions

We have provided novel etymologies for six Romanian substrate words with Albanian comparanda: *berc*, *driză*, *groapă*, *mărar*, *rangă* and *urdă*. The first two words are identifiable as Albanian loans. In *groapă*, Romanian preserves a relic from a minor Paleo-Balkan language — Paeonian or Brygian — likely acquired via Albanian. A chain borrowing scenario explains *mărar*, a word of Ancient Greek origin mediated to Romanian by Thracian and Albanian. The preserved /an/ in *rangă*, a Thracian loanword, imposes a 6th-century AD terminus post quem for the end of Thracian-Balkan Latin contacts. Finally, *urdă* is also analyzable as a Thracian loan, with the Romanian word being the most probable source of Alb. *urdhë*.

In addition to providing new solutions to problems of Romanian etymology, our work provides insights on extinct Paleo-Balkan languages — most notably, evidencing a PIE **w* > Thr. $\emptyset$ / $_$ *{*_r, *_ṛ}* shift. This result illustrates how data from surviving languages of Southeastern Europe can reveal phonological features of the sparsely attested extinct languages once spoken in the same area. Romanian, in particular, holds great promise in this regard. While we have chosen to limit the scope of this paper to substrate words with Albanian comparanda, Romanian may offer a considerably richer substrate corpus, given that the language reportedly contains over 1,200 words of unknown etymology according to recent counts (Paliga 2022: 235). In sum, we maintain that treating the Romanian substrate as a composite of distinct layers not only offers the most suitable framework for advancing its understanding, but also helps uncover features of Paleo-Balkan languages that would be inaccessible through the ancient attested data alone.

Abbreviations and symbols

AGr. — Ancient Greek; Alb. — Albanian; AMac. — Ancient Macedonian; ARo. — Aromanian; C — consonant; Gr. — Greek; H — Proto-Indo-European laryngeal; Hun. — Hungarian; Lat. — Latin; OCS — Old Church Slavonic; OHG — Old High German; PGm. — Proto-Germanic; PIE — Proto-Indo-European; PSl. — Proto-Slavic; Rom. — Romanian; T — plosive; Thr. — Thracian; V — vowel.

³⁶ An example of this influence is the Rom. *-esc-* suffix, one of the most important Romanian suffixes, which continues Thr. *-isk-* (Graur 1927: 549–551).

References

- Ajeti, Idriz. 2017. Formimi i fjalëve dhe pasurimi i gjuhës shqipe. In: Rexhep Ismajli (ed.). *Studime për gjuhën shqipe*: 511–539. Prishtinë: Akademia e Shkencave dhe e Arteve e Kosovës.
- Arap, Ina. 2012. Albanisch-rumänische Parallelen im Bereich der Wortbildung: Die Suffixe. In: Wolfgang Dahmen, Günter Holtus, Johannes Kramer, Michael Metzeltin, Wolfgang Schweickard, Otto Winkelmann (eds.). *Südosteuropäische Romania. Siedlungs-/Migrationsgeschichte und Sprachtypologie. Romanistisches Kolloquium XXV*: 123–152. Tübingen: Narr Francke Attempto Verlag.
- Barić, Henrik. 1955. *Hymje në historin e gjuhës shqipe*. Prishtinë: Shtëpia botuese “Mustafa Bakija”.
- Beekes, Robert S. P. 2010. *Etymological dictionary of Greek*. 2 vols. Leiden / Boston: Brill.
- Blažek, Václav. 2006. Bohumil Vykypěl, Studie k šlechtickým titulům v germánských, slovanských a baltských jazycích. *Baltistica* 41(3): 511–518. DOI: <https://doi.org/10.15388/baltistica.41.3.1363>
- Boerescu, Pârvu. 2018. *Elemente de substrat (autohtone) ale limbii române: — Compendiu —*. București: Editura Academiei Române.
- Bogrea, Vasile. 1921. Elemente latine. *Dacoromania. Buletinul „Muzeului Limbei Române”* 1: 257–260.
- Brâncuș, Grigore. 1983. *Vocabularul autohton al limbii române*. București: Editura Științifică și Enciclopedică.
- Brâncuș, Grigore. 2005. *Introduce în istoria limbii române*. București: Editura Fundației “România de Măine”.
- Brâncuș, Grigore. 2013. Problemele substratului limbii române. *Limba Română* 4: 417–423.
- Brâncuș, Grigore. 2022. Definirea substratului limbii române. *Limba Română* 32(3–4): 46–57.
- Çabej, Eqrem. 1964. Studime rreth etimologjisë së gjuhës Shqipe XII: K. *Studime Filologjike* 18(1): 61–98.
- Candrea, Ion Aurel, Ovid Densusianu. 1907. *Dicționarul etimologic al limbii române. Elementele latine. 2. Cere — fn*. București: Librăria Socec & Comp.
- Capidan, Theodor. 1922. Raporturile albano-române. *Dacoromania. Buletinul „Muzeului Limbei Române”* 2: 444–553.
- Cihac, Alexandru. 1879. *Dictionnaire d'étymologie daco-romane. Éléments slaves, magyars, turcs, grecs-moderne et albanais*. Francfort: Ludolphe St-Goar.
- Ciorănescu, Alexandru. 2001. *Dicționarul etimologic al limbii române*. București: Saeculum I. O.
- de Vaan, Michiel. 2008. *Etymological dictionary of Latin and the other Italic languages*. Leiden / Boston: Brill.
- DELR = Academia Română. 2025. *Dicționarul etimologic al limbii române*. Available online at: <https://delr.lingv.ro/>
- Demiraj, Bardhyl. 1997. *Albanische Etymologien (Untersuchungen zum albanischen Erbwortschatz)*. Amsterdam / Atlanta: Rodopi.
- Derksen, Rick. 2008. *Etymological dictionary of the Slavic inherited lexicon*. Leiden / Boston: Brill.
- Diculescu, Constantin C. 1927. Elemente vechi grecești din limba română. *Dacoromania. Buletinul „Muzeului Limbei Române”* 4(2): 393–516.
- DLR = Academia Română. 1913–1949. *Dicționarul limbii române. 2 vols*. București: Librăriile Socec & Comp. și C. Sfetea / Imprimeria Națională / Tipografia Ziarului „Universul”.
- Duridanov, Ivan. 1985. *Die Sprache der Thraker: Vom Verfasser übersetzt und wesentlich erweitert*. Sofia: Hieronymus Verlag Neuried.
- Duridanov, Ivan. 1993. Brygische Spuren im thrakischen Sprachgebiet. *ORPHEUS. Journal of Indo-European and Thracian Studies* 3: 63–68.
- Duridanov, Ivan. 1995. Thrakisch — Dakisch — Pelasgisch. In: Michaela Ofitsch, Christian Zinko (eds.). *Studia onomastica et indogermanica: Festschrift für Fritz Lochner von Hüttenbach zum 65*: 45–57. Graz: Leykam Verlagsgesellschaft.
- Duridanov, Ivan. 2000. Die Rolle der Paläobalkanistik für die Südosteuropa-Linguistik. In: Uwe Hinrichs, Uwe Büttner (eds.). *Die Südosteuropa-Wissenschaften im neuen Jahrhundert. Akten der Tagung vom 16.–19. Oktober 1999 an der Universität Leipzig*: 23–32. Wiesbaden: Harrassowitz Verlag.
- Friedman, Victor A., Brian D. Joseph. 2025. *The Balkan languages*. Cambridge University Press.
- Georgiev, Vladimir I. 1965. Le dace comme substrat de la langue roumaine. *Revue roumaine de linguistique* 10(1–3): 75–80.
- Georgiev, Vladimir I. 1981. *Introduction to the history of the Indo-European languages*. Sofia: Publishing House of the Bulgarian Academy of Sciences.
- Georgiev, Vladimir I. 1983. Thrakisch und Dakisch. In: Wolfgang Haase (ed.). *Aufstieg und Niedergang der römischen Welt: Geschichte und Kultur Roms im Spiegel der neueren Forschung: II*: 1148–1194. Berlin / New York: Walter de Gruyter.

- Graur, Alexandru. 1927. Le suffixe roumain *-escu* et le suffixe thrace *-isk*. *Romania* 53: 539–552.
- Greenberg, Marc L. 2017. Slavic. In: Mate Kapović (ed.). *The Indo-European languages. Second edition*: 519–551. London / New York: Routledge.
- Hamp, Eric P. 1976. Some Romanian areal etymologies. *Studii și cercetări lingvistice* 27(1): 33–36.
- Hill, Peter M. 2011. Language contact in South-Eastern Europe. In: Bernd Kortmann, Johan van der Auwera (eds.). *The languages and linguistics of Europe. A comprehensive guide*: 409–428. Berlin / Boston: Mouton De Gruyter.
- Hubschmid, Johannes. 1983. Zum Substrat und zur Vorgeschichte des Rumänischen (Probleme der Balkanlinguistik). *Zeitschrift für romanische Philologie* 99(5–6): 497–511. DOI: <https://doi.org/10.1515/zrph.1983.99.5-6.497>
- Huld, Martin E. 1986. Accentual stratification of Ancient Greek loanwords in Albanian. *Zeitschrift für vergleichende Sprachforschung* 99(2): 245–253.
- Hyllested, Adam. 2016. Proto-Indo-European reconstruction and Albanian phonotactics. In: Stephanie W. Jamison, H. Craig Melchert, Brent Vine, Angelo Mercado (eds.). *Proceedings of the 26th Annual UCLA Indo-European Conference: Los Angeles, October 24th and 25th, 2014*: 63–81. Bremen: Hempen.
- Ivănescu, Gheorghe. 1980. *Istoria limbii române*. Iași: Junimea.
- Kálmán, Béla. 1972. Hungarian historical phonology. In: Loránd Benkő, Samu Imre (eds.). *The Hungarian language*: 49–84. The Hague / Paris: Mouton.
- Kaluzhskaya, Irina A. 1993. Paleobalkan linguistic heritage and the languages of Balkan-Carpathian area. *OR-PHEUS. Journal of Indo-European and Thracian Studies* 3: 69–72.
- Katičić, Radoslav. 1976. *Ancient languages of the Balkans: Part one*. The Hague / Paris: Mouton.
- Khoshsirat, Zia, Andrew Miles Byrd. 2023. The Indo-Iranian labial-extended causative suffix. *Indo-European Linguistics* 11(1): 64–95. DOI: <https://doi.org/10.1163/22125892-bja10025>
- Klotz, Emanuel. 2017. *Urslawisches Wörterbuch*. Wien: Facultas.
- Ködderitzsch, Rolf. 2001. Albanisch und Thrakisch. *Studia Albanica* 34(1): 79–87.
- Kroonen, Guus. 2013. *Etymological dictionary of Proto-Germanic*. Leiden / Boston: Brill.
- Leschber, Corinna. 2021. Urslavisches Lexikon in Rumänischen. In: Anna Jouravel, Audrey Mathys (eds.). *Wort- und Formenvielfalt: Festschrift für Christoph Koch zum 80. Geburtstag*: 207–218. Berlin: Peter Lang.
- LIV = Rix, Helmut, Martin Joachim Kümmel, Thomas Zehnder, Reiner Lipp, Brigitte Schirmer. 2001. *Lexikon der indogermanischen Verben*. Wiesbaden: Dr. Ludwig Reichert Verlag.
- Lubotsky, Alexander, Sergei Starostin. 2003. Turkic and Chinese loan words in Tocharian. In: Brigitte L. M. Bauer, Georges-Jean Pinault (eds.). *Language in time and space. A Festschrift for Werner Winter on the occasion of his 80th birthday*: 257–270. Berlin / New York: Mouton de Gruyter.
- Maiden, Martin, Adina Dragomirescu, Oana Uță Bărbulescu, Gabriela Pană Dindelegan, Rodica Zafiu. 2020. *The Oxford history of Romanian morphology*. New York: Oxford University Press.
- Matzinger, Joachim. 2006. *Der altalbanische Text Mbsuame e kreshtere (Dottrina cristiana) des Leke Matrenga von 1592. Eine Einführung in die albanische Sprachwissenschaft*. Dettelbach: Verlag J. H. Röhl GmbH.
- Matzinger, Joachim. 2017. L'elemento latino della lingua albanese — un impatto della Via Egnatia? *Palaver* 6(2): 29–60.
- Matzinger, Joachim. 2018. 98. The lexicon of Albanian. In: Jared Klein, Brian D. Joseph, Matthias Fritz, Mark Wente (eds.). *Handbook of comparative and historical Indo-European linguistics. HSK 41.3: 1788–1800*. Berlin / Boston: Walter de Gruyter GmbH.
- MDA = Academia Română. 2001–2003. *Micul dicționar academic*. 4 vols. București: Editura Univers Enciclopedic.
- Meyer, Gustav. 1891. *Etymologisches Wörterbuch der albanesischen Sprache*. Strassburg: Verlag von Karl J. Trübner.
- Mihaylova, Bilyana, Albena Mircheva. 2022. The Thracian glosses revisited. In: Peter Delev et al. (eds.). *Ancient Thrace: Myth and reality. The proceedings of the Thirteenth International Congress of Thracology. Kazanlak, September 3 — 7, 2017*. Vol. 2: 209–217. Sofia: St. Kliment Ohridski University Press.
- Miklosich, Franz. 1861. *Die slavischen Elemente im Rumunischen*. Wien: Kaiserlich-königliche Hof- und Staatsdruckerei.
- Mircheva, Albena, Bilyana Mihaylova. 2020. The reflexes of the Indo-European laryngeals in the Paleo-Balkan languages. *Linguistique balkanique* 59(1): 46–61.
- Nagy, Levente. 2003. Balcanistica hungarica rediviva. *Zeitschrift für Balkanologie* 39(2): 191–200.
- Neroznak, V. P. 1978. *Paleobalkanskije jazyki*. Moskva: Nauka.
- Orel, Vladimir E. 1998. *Albanian etymological dictionary*. Leiden / Boston / Köln: Brill.
- Paliga, Sorin. 2022. Thracian v. Dacian (Daco-Moesian). How many 'Thracian languages' were there spoken in the Antiquity? In: Peter Delev et al. (eds.). *Ancient Thrace: myth and reality. The proceedings of the thirteenth interna-*

- tional congress of Thracology. Kazanlak, September 3 — 7, 2017. Vol. 2: 231–238. Sofia: St. Kliment Ohridski University Press.
- Paliga, Sorin. 2024. *An etymological dictionary of the Romanian language*. New York: Peter Lang.
- Petrovici, Emil. 1948. Note slavo-române. III. *Dacoromania*. *Buletinul „Muzeului Limbei Române”* 11: 184–193.
- Philippide, Alexandru. 1927. *Originea românilor: Volumul al doilea: Ce spun limbile română și albaneză*. Iași: Tipografia „Viața Românească” S. A.
- Poghirc, Cicerone. 1969. Influența autohtonă. In: Ion Coteanu, Gh. Bolocan, Matilda Caragiu Marioțeanu (eds.). *Istoria limbii române*. 2: 313–365. București: Editura Academiei Republicii Populare Române.
- Pokorny, Julius. 1959. *Indogermanisches etymologisches Wörterbuch*. Vol. 1. Bern / München: Francke Verlag.
- Pușcariu, Sextil, Matteo Bartoli, Alois Belulovici, Arthur Byhan. 1926. *Studii istroromâne*. 2. *Introducere. Gramatică. Caracterizarea dialectului istroromân*. București: Cultura Națională.
- Ramstedt, Gustaf John. 1951. Alte türkische und mongolische Titel. *Journal de la Société Finno-ougrienne* 105(2): 59–82.
- Rusakov, Alexander. 2017. Albanian. In: Mate Kapović (ed.). *The Indo-European languages. Second edition*: 552–608. London / New York: Routledge.
- Rusakov, Alexander. 2018. 94. The documentation of Albanian. In: Jared Klein, Brian D. Joseph, Matthias Fritz, Mark Wenthe (eds.). *Handbook of comparative and historical Indo-European linguistics. HSK 41.3*: 1716–1731. Berlin / Boston: Walter de Gruyter GmbH.
- Russu, Ioan I. 1981. *Etnogeneza românilor. Fondul autohton traco-dacic și componenta latino-romană*. București: Editura Științifică și Enciclopedică.
- Șăineanu, Lazăr. 1929. *Dicționar universal al limbei române*. București: Editura „Scrisul Românesc”.
- Saramandu, Nicolae. 1987. Concordanțe etimologice și lexicale româno-albaneze (bearcă, buș, ‘pumn’, urdori, cocobarză, corcodușă). *Studii și cercetări lingvistice* 38(2): 126–141.
- Schade, Oskar. 1872–1882. *Altdeutsches Wörterbuch*: 1. Halle: Verlag der Buchhandlung des Waisenhauses zu Halle.
- Schumacher, Stefan. 2009. Lehnbeziehungen zwischen Protoalbanisch und balkanischem Latein bzw. Romanisch. In: Oliver Jens Schmitt, Eva Anne Frantz (eds.). *Albanische Geschichte: Stand und Perspektiven der Forschung*: 37–59. München: Oldenbourg Verlag.
- Schumacher, Stefan. 2013. Historische Phonologie. In: Stefan Schumacher, Joachim Matzinger (eds.). *Die Verben des Altalbanischen: Belegwörterbuch, Vorgeschichte und Etymologie*. Wiesbaden: Harrassowitz Verlag.
- Schütz, István. 2002. *Fehér foltok a Balkánon: Bevezetés az albanológiába és a balkanisztikába*. Budapest: Balassi Kiadó.
- Scriban, August. 1939. *Dicționarul limbii românești. Etimologii, înțelesuri, exemple, citațiuni, arhaisme, neologisme, provincialisme*. Iași: Institutul de Arte Grafice “Presa Bună”.
- Slușanschi, Dan. 1981. Substratul daco-moesian al limbii române. *Limba Română* 30(4): 339–342.
- Sutherland, Mhairi A., Cassandra B. Tucker. 2011. The long and short of it: A review of tail docking in farm animals. *Applied Animal Behaviour Science* 135(2): 179–191. DOI: <https://doi.org/10.1016/j.applanim.2011.10.015>
- Tiktin, Heimann Hariton. 1903–1925. *Rumänisch-deutsches Wörterbuch*. 3 vols. București: Imprimeria Statului.
- Ungureanu, Dan. 2016. *Româna și dialectele italiene*. București: Editura Academiei Române.
- Vinereanu, Mihai. 2023. *Dicționarul etimologic al limbii române*. 2 vols. București: Editura Uranus.
- Witczak, Krzysztof Tomasz. 1995. Were the Bessans ancestors of the Albanians? A new opinion on the ethnogenesis of the Albanian nation. *Thraco-dacica* 16(1–2): 309–312.
- Yanakieva, Svetlana. 2009. *Trakiyska hidronimiya*. Sofiya: Akademichno izdatelstvo “Prof. Marin Drinov”.
- Yanakieva, Svetlana. 2018. The Thracian language. *ORPHEUS. Journal of Indo-European and Thracian Studies* 25: 26–68.
- Zaicz, Gábor. 2006. *Etimológiai szótár. Magyar szavak és toldalékok eredete*. Budapest: TINTA Könyvkiadó.
- Zhan, Zhang. 2014. Secular Khotanese documents and the administrative system in Khotan. *Bulletin of the Asia Institute* 28: 57–98.

Андрей Ф. Стойка, Сюэ К. Р. Стойка. О нескольких румынских субстратных словах с албанскими соответствиями: этимология и импликации

В статье предлагается новый взгляд на этимологии некоторых румынских субстратных слов, имеющих близкие албанские соответствия, с упором на определение языка-источника для каждого слова. Выявлены два новых албанских заимствования: *berc* ‘с ко-

ротким хвостом; бесхвостый' и *driză* 'мохнатый пырей', тогда как *rangă* 'лом' и *urdă* 'вид сыра, похожий на рикотту' определяются как фракийские заимствования. Помимо этого, для *groapă* 'яма' (< **g^hrob^h-eh₂*) предполагается происхождение от субстрата, сохранившего праиндоевропейское **o* (пеонийского или бригийского). Наконец, *mărar* 'укроп' выводится в конечном счете из древнегреческого **maráthrion* (дими́нутив от *máraithron* 'фенхель'), через посредничество как фракийского, так и албанского.

Ключевые слова: румынский язык; румынская этимология; палеобалканские языки; албанский язык; фракийский язык.