

A FEW NOTES ON PERSIAN LOANWORDS IN ROMANI

ALIN-GABRIEL BUCĂ*

Abstract. The Iranian lexicon in Romani has not been studied in depth from a more comprehensive perspective and multiple challenges arise when trying to identify direct Persian loans. Indo-Iranian and Indo-Aryan share a substantial number of cognates and most lexicographic works with etymological glosses generally give precedence to the latter. Mediation via Turkish or another Balkan L2 needs to be evaluated before weighing the likelihood of a direct Persian origin, and a fine-tuned phonological analysis needs to be carried out. In some cases, comparison with the Lomavren corpus can provide valuable clues for the elucidation of murkier etymological pathways. A sound survey of dialectal isoglosses is also required for a clearer picture of the underlying lexical borrowing processes. The paper provides a set of sample cases glossed using this approach, which could form the basis of more extensive future research.

Keywords: Romani; Persian; Turkish; Lomavren; historical phonology; lexicographic sources

1. Introduction

This paper aims to sketch out a more comprehensive and nuanced approach towards the analysis of Persian loans in Romani. Drawing on existing studies on Farsisms in Turkish (Tietze & Lazard 1967, Pomorska 2013, Stachowski K. 2020), a preliminary step is to rule out Turkish mediation on phonetic, historical or semantic grounds. Whenever possible, this should be complemented by a dialect-oriented analysis, so as to appraise the relative distribution, relevance and stability of certain words within Romani at large, for instance Turkish-heavy dialects such as Spoitori are certainly outliers (see also Friedman 2020). Equally important, the presence of the voiceless velar fricative [x] can act as a

* University of Bucharest, “Limbi și Identități Culturale” Doctoral School, alin.buca@s.unibuc.ro

diachronic marker and, when in auslaut, it almost unequivocally points towards a Persian origin. In the case of some cognates, Lomavren could act, albeit with limited applicability, as a ‘control’ group against which larger loan tendencies or individual occurrences could be tested. After reviewing the available Romani glossaries and lexicographic resources, we have selected a representative set of sample cases which illustrate the aforementioned methodological steps at work. We are confident that this approach can be further refined and applied to a larger corpus.

2. Romani, Persian, Turkish: intertwined contacts

Romani is a diasporic New Indo-Aryan language whose diachronic trajectory is anchored between the South-Asian and the Balkan linguistic areas. The intermediate leg of this journey, following the departure of the proto-Roma from the Indian subcontinent in the 11th century¹ and their arrival in Persian-speaking territories, overlaps with the influx of Turkic peoples from Central Asia, which marked the beginning of an intense process of linguistic diffusion that substantially impacted Persian² and other Iranian languages (Windfuhr 2009). This was undoubtedly a two-way convergence, as Turkic dialects also became Iranianized to a certain extent, at times to the point of symbiosis (Johanson 2010). On the other hand, it must be noted that both Persian and Turkish functioned as *linguae francae* throughout the Persian Empire and even beyond its borders, along commercial routes, and we may indeed talk about a Turkish-Persian continuum. This particular context, coupled with the later, more intense exposure of proto-Romani to Turkish in Anatolia and the Balkans, complicates any attempt to isolate unmediated Persian loanwords in the Romani lexicon. A maximalist mediated borrowing chain would be Persian > [Turkish] > [Balkan L2] > Romani³, so several possible intermediate steps must be ruled out before determining a direct acquisition path.

¹ This was arguably marked by the capture of Kannauj in 1018 by the sultan Maḥmūd of Ghaznī (Courthiade 2016).

² Conventional historical stages: Middle Persian (4th century BC – 7th century AD); Early New Persian (7th century – 13th century); Classical New Persian (13th century – 18th century); Modern and Contemporary Persian (18th century – present) (Maggi & Orsatti 2018).

³ Stachowski M. (2019) states, rather surprisingly, that “Another source of Turkic words in the Balkan languages was the language of the Gypsies”, basing this

3. Research background and method

In the preliminary stage of our research, we reviewed the word lists available in the literature and selected a number of entries (*baxt*, *bezex*, *buzex*, *buzno*, *xulaj*, *pendex*, *taxtaj*, *zet*) which best lend themselves to a critical and/or more complex analysis and form the majority of the sample cases discussed herein. Secondly, we consulted the main Romani lexicographic works containing etymological annotations, which are listed *ad litteram* in each case. Finally, we referenced the relevant resources on Lomavren (Finck 1905, Voskanian 2002, Scala 2014), aiming to use the Lomavren lexicon, where possible, as a ‘control’ corpus.

3.1. Existing glossaries

Several glossaries of putative Iranisms in Romani have been published up to date. Friedman (1989) attempted to define the position of Turkisms in Romani and carried out a detailed analysis of the Agia-Varvara dialect, dividing the identified corpus into three categories: Arabisms (undoubtedly mediated by Turkish and/or Persian), Turkisms, and Farsisms. However, an in-depth etymological analysis is not provided, and the diachronic component is neglected. As a result, Friedman’s list includes a sizable number of recent Balkan borrowings. Boretzky & Igla (1994) provide a relatively detailed inventory of Iranian and Armenian loanwords in Romani, along with pertinent etymological glosses. Interestingly though, in Boretzky (2012) we find a mixed list of Iranisms and Armenisms, which only comes to show the challenge of distinguishing the primary source language, especially in the case of Armenian, which contains many Iranisms itself. Hanckock (1995) embarked on a more ambitious endeavor, taking an inventory of Iranisms in Romani, Lomavren, and Domari. Unfortunately, the Romani glossary includes quite a few recent, most probably Turkish-mediated loanwords.

assumption on the claim made by some historians that “the earliest groups of Gypsies in Romania were brought in the 13th century as slaves of invading Tatars”. However, this historical hypothesis has been refuted convincingly by Achim (2004), and Stachowski fails to provide further linguistic evidence.

3.2. *Lexicographic resources*

In this paper, we have referenced the main Romani dictionaries that contain etymological annotations: Sampson (1926), while ostensibly covering only Welsh Romani, is replete with extremely detailed and valuable information, pinpointing cognates and equivalents in various Indo-Aryan and Indo-Iranian languages; Boretzky & Igla (1994) cover a sizeable number of dialects and also provide word lists classified by etyma, as well as a grammar compendium; Mānušs *et al.* (1997), even though focusing on Baltic Romani, puts forth interesting etymologies applicable to the core Romani lexicon at large; Vekardi (2000) covers the Romani dialects spoken in Hungary and provides interesting etymological suggestions; Calvet (2021) focuses on the Kalderash dialect and records multiple etymologies and cognates for most entries; Mégret (2016) seems to draw substantial inspiration from other sources and oftentimes falls into fanciful speculations.

Other Romani dictionaries referenced in this paper include Courthiade (2009), Lee (2010), Sarău (2006, 2012), and Tahirović-Sijerčić (2013), which do not include etymological entries. For cross-dialectal distribution and isogloss analysis, we have conducted queries in the ROMLEX database.

Unless otherwise indicated, for all Sanskrit, Persian and Turkish references we have consulted Apte (1993), Aryanpur (2017) and Parker (2008), respectively.

3.3. *Lomavren: viable ‘control’ corpus?*

Sampson (1926) argued for a sub-group affiliation of Lomavren (documented in Armenia), Domari (spoken in Syria, Palestine, Egypt), and Romani, which would have branched out from a common NIA ancestor. However, Turner (1926) offered sound counterarguments, and currently the prevailing opinion remains that these are different, though similar, languages.

The linguistic affinity between Romani and Domari (and, as far as documented, Lomavren) might therefore be accounted for in terms of their shared ancient origin and subsequent similar social and geographical

history, rather than as a token of continuous genetic ties in the form of a linguistic sub-branch within the Indo-Aryan languages.

Scala (2014) considers Lomavren an “intertwined language”, characterized by a split between grammar and lexicon. Notwithstanding its Armenian-derived grammar, Lomavren still retains a large Indo-Aryan lexicon and was similarly subject to contact with Persian (as evidenced by the available corpus, though in many instances Armenian-mediation must have occurred). As such, a contrastive analysis of cognates shared with Romani can be helpful.

4. Phonological remarks

Historical phonology allows us to distinguish certain inflection points with respect to the relationship between Romani and various contact languages, and in terms of its internal convergences and divergences at a dialectal level. Sampson (1926) lists a series of phonetic changes claimed to have occurred outside the Indian subcontinent, in Persia, Armenia and Syria. Most notable is the adoption into proto-Romani of the voiceless velar fricative [x], which is absent in Indo-Aryan languages, both in Persian loanwords and in some inherited words where it replaces the initial voiceless aspirated velar stop [kh] (loss of aspiration + fricativization) (e.g. *xal* ‘to eat’ vs. Skt. खाद् *khād*, Prākr. खाइ *khāi*, Hi. खाना *khānā*, Pers. خوردن *xordan*). As such, [x] represents an important indicator which generally helps to distinguish between the old lexical stock and more recent acquisitions. In some cases, as detailed below, it may lend more weight to the likelihood of a Persian etymology.

4.1. The voiceless velar fricative in Persian

While Middle Persian had both [x] and the labiovelar fricative [x^w], the two phonemes are gradually fused in Early New Persian [x^w] > [x]. In Modern Persian, [x] is characterized by a high degree of flexibility and fluctuation, having no fewer than three positional variants (a velar, a post-velar, and a uvular allophone) which cover the entire velar-uvular

area. At the same time, there is no attested convergence with the glottal fricative [h]. The voiceless uvular stop [q] has an even wider allophonic spectrum, comprising both stop ([q], [g]) and fricative allophones ([χ], [ʁ]). Pisowicz (1985) points out that in the colloquial register the opposition [q]-[x] is oftentimes neutralized by spirantization [q] > [x]. This is made possible by the allophonic overlap of the two phonemes. Pisowicz goes on to analyze the pair وقت *vaqt* [vaxt] 'time' - بخت *baxt* [baxt] 'luck', concluding that the two can be differentiated because [x] is interchangeable with an occlusive allophone of [q] in the former, but remains etymological and thus non-switchable in the latter [x].

4.2. *The voiceless velar fricative in Romani*

In Romani, [x] generally appears in words that entered the language prior to the contact with Byzantine Greek, and [h] in more recent acquisitions from other contact languages (Romanian, Hungarian, Turkish, etc.). Naturally, there are exceptions, such as *xoli* 'bile, anger', *xarkum(a)* 'copper', *oxto* 'eight', which are all Greek borrowings. Regarding the [x]-[h] dynamic at dialectal and sub-dialectal level, the picture is rather heterogeneous: some dialects preserve the distinction between the two, but [h] and [x] can be interchangeable in some words; in other cases, the phonemes are assimilated, either [x] > [h], more frequently, or [h] > [x], less so. As in Persian, [x] possesses a relatively wide phonemic spectrum. For instance, voicing ([x] > [ɣ]) is seen in the Polish Xaladitka dialect, while in other dialects we encounter uvular variants, e.g. the voiceless uvular fricative [χ] in the Greek Xoraxane dialect, the voiced uvular fricative [ʁ] in Lithuania (Baló 2020).

5. Sample cases

In the following we will discuss a few relevant sample cases. Each item is glossed using the available Romani lexicographic resources which contain relevant etymological entries (Sampson 1926, Boretzky & Igla 1994, Mănuș et al. 1997, Vekardi 2000, Mégret 2016) and then discussed briefly.

baxt f. 'luck', 'happiness'

Sampson: Generally regarded as an Iranian loan-word from Pers., Afgh., Kurd. *bakht* 'fortune, 'happiness'. The word may however be original; cp. Skr. *bhaga*, *bhāgya* 'lot', 'good fortune', 'luck', 'happiness', *bhāgtā* 'happiness' from √*bhaj* 'to allot', Hind. *bhāg*. From *bhaga* or *bhāgtā* Gyp. *baxt* may have arisen through the forms **baghta*, **baght*, *bakht*, *baxt*. **Boretzky & Igla:** < pers. *baxt*. **Vekardi:** < pers. *baxt*. **Mānušs:** < viduspers.⁴ un senarm. ⁵ *baxt* 'happiness, luck; fate, destiny'. **Calvet:** *bax*; T.c., *baxt*; person, بخت (*baxt*), s., 'chance, fortune, sort'; arm., բախտ (*baxt*), s., 'bonheur, chance'. **Mégret:** *bax* < srk. भग *bhaga* 'chance', a subi ultérieurement l'influence du pers. *baxt* 'chance', mais cela ne justifie pas l'orthographe *baxt* au lieu de *bax*.

An Indo-Aryan etymology is possible (metathesis of aspiration>**bagha* + devoicing > **bakha*), but quite unlikely, as fricativization of velar stops ([k] > [x]) is a rare and irregular phenomenon in Romani (Matras, 2004). On the other hand, the Turkish form is *baht*, a Persian loan which displays a regular sound change, Ar./Pers. [x] > Tur. [h] (Stachowski K. 2020), which rules out Turkish mediation. In light of these considerations, Romani *baxt* is almost certainly borrowed *tel quel* from Persian (بخت *baxt*). It bears stressing that Mégret's argument is certainly unsound, as the variant *bax* is common in dialects which display final stop or cluster reduction in the nominative singular (for instance *gra'/gras* < *grast*). However, the 'phantom' stop/cluster can be readily observed in oblique and plural forms.

bar f. n. 'fence, enclosure', 'court-yard'

Sampson: *bār̥*⁶. Skr. *vāṭa* (m.) 'enclosure', *vāṭikā* (f.) 'garden', Hind. *bārī* 'enclosure', 'garden', 'orchard', *bār* 'fence', 'hedge'. **Boretzky & Igla:** < pk. *vāḍa*- < ai. *vāṭa*-, *vāṭi*. **Vekardi:** <skr. *vāṭa*?; <Pers. *bāgh*?. **Mānušs:** *bār*. < SIA⁷:

⁴ Middle Persian

⁵ Old Armenian

⁶ In Sampson (1926) *ā* indicates low-back-round *ɒ*. The minimal pair counterpart *barr* 'stone' is noted as *bār*.

⁷ Old Indo-Aryan (OIA).

skt. *vāṭī* ‘enclosure; court, yard, court-yard; garden’. **Calvet**: Skr. **वाट** (*vāṭa-*), s.m., ‘enclose, jardin’, **वाटि** (*vāṭi-*), s.f., ‘enclose’; hin. **बाड़** (*bār*) s.f., ‘clôture, barrière’, **बाड़ी** (*bārī*) s.f., ‘jardin, verger, enclose’. **Mégret**: <skr. **वाटी**-*vāṭi* ‘jardin potager’, hi. **बाड़ी** *bārī* ‘petit verger, enclos’, pers. *bagh* ‘jardin’.

Even though most lexicographic sources give credence to an Indo-Aryan origin, we believe there are arguments supporting a Persian filiation. باغ *bāq* [bāʁ] ‘garden’ is a native Persian word. Phonetically, in Modern Persian, the letter <غ> [qeyn] corresponds, along with the letter <ق> [qāf], to the same phoneme [q], regardless of position. However, in Classical New Persian, same as in Arabic, these letters designated two separate phonemes: the voice uvular fricative [ʁ] <غ> and the voice uvular stop [q] <ق>, respectively (Pisowicz 1985). The distinction remains at the allophonic level, where we find both [bāq] (formal pronunciation) and [bāʁ] (in the colloquial register, but also in intervocalic position, e.g. باغ و *va* [o] > [bāʁo] ‘the garden and...’ (Pisowicz 1985). Also, in the proximity of voiceless consonants, we encounter another allophone, the voiceless velar fricative [x]: باغچه *bāqčē* [baxtʃe] ‘small garden’ (Ghavami 2018). A putative borrowing into proto-Rromani would have brought along the distinct phoneme [ʁ] and its series of allophones, transformed subsequently through lenition into the alveolar vibrant [r].

A Turkish-mediated origin is rather unlikely, especially on semantic grounds: bağ entered Turkish at an early date, with the meaning ‘vineyard’, for which Turkish lacked a native equivalent (Clauson 1972, Pomorska 2013), while for ‘garden’ the Persian diminutive was adopted باغچه *bāqčē* > *bahça*, *bağça*, *bağçe* (Pomorska 2013), which became in Modern Turkish *bahçe*, through debuccalization of the internal consonant.

Another interesting aspect is the dynamic of *bar* and its minimal pair *barr* m.n. ‘stone’. *Barr* is an extremely stable native word attested in most Romani dialects (< OIA **वटः** *vaṭaḥ* ‘cowrie’, ‘small globular object’, ‘pellet’, Prākr. *vāḍa*) (Sampson 1926) which displays regular sound changes: internal [ṭ] (via MIA [d] succeeded by [r] and preserved in some dialects as uvular [ʁ] or long trill [rr] + shift from labial fricative to stop in initial position (Matras 2004). In contrast, an RMS database query shows that *bar* appears in a significantly smaller number of dialects,

indicating that it might very well be a newer, less stable acquisition. In many dialects we encounter recent L2 loanwords (Slavic – *dvoro*, *udvara*, *sado*, *dvora*, *pridvora*, *greda*, *oboros*, *boro*, *gradina*; Turkish – *baxčik*, *bavča*, *baxčas*; Romanian – *kimpo*, *batatura*; other – *veranda*, *džardino*, *seta*, *munda*, etc.) or other native equivalents ((a)vri ‘outside’, *avlin*, *avlija*). In cases where the pair is attested, we encounter near-homophones where the liquid quality remains the distinctive feature [rr/ř/r] vs. [r], perfect homophones differentiated solely by the nominal gender, or fully merged homophones. Interestingly, a similar homophonous convergence phenomenon occurs in Lomavren, albeit involving different etyma, *bar* ‘door’ < MIA वारो *varo* OIA द्वार and *bar* ‘against’ < *barbar* (< Pers. بر *bar*).

bezex m. n. ‘sin’

Sampson: –. **Boretzky & Igla:** < pers. *bazah*. **Vekardi:** < pers. *bazah*. **Mānuš:** < viduspers.⁸ *bazág* ‘sin’. **Calvet:** persan, بهزه (*beze*), s., ‘péché, faute’. **Mégret:** *bezex*, *bizex* < pers. *beze*.

Evidently, this is a direct Persian loan بهزه [bezeh] which is also documented in Lomavren *bazax* (Finck, 1905). Voskanian (2002) argues that the Lomavren word originates in the Classical New Persian, resulting from the velarization of the glottal fricative [h]. Hancock (1995) mentions Kurdish and Turkish counterparts (*beze*). However, we could not identify the word in the referenced Turkish and Kurdish dictionaries. A Turkish mediation is certainly far from likely, as Pers. [h] > Tur. Ø is a rare phenomenon, [h] > [h] being generally the case (Stachowski K. 2020).

buzex f. n. ‘spur’, ‘hoof’

Sampson: *būzex*. According to Pott by metath. fr. Pers. *sbukh* ‘a prick’, *sbukhtan*, *spuzidan* ‘to stab’, ‘to pierce’. **Boretzky & Igla:** vgl. pers. *spuzīdan* ‘stechen, durchbohren’. **Vekardi:** –. **Mānuš:** *buzéx* < arm. *bazóg* ‘poking, butting, pricking; thrusting, pushing, urging on’. **Calvet:**

⁸ Old Persian.

Métathèse du persan سپوختن (*sepuxtān*), سپوز (*sepuz*), v., ‘enfoncer par force, trouer’. **Mégret**: < pers. *sepuxtān* ‘percer’.

Phonologically, the Armenian etymology adduced by Mānušs is more convincing and straightforward than سپوختن *sepuxtān*. This word is seemingly not attested in Lomavren, though more thorough research of the corpus could prove otherwise. Further investigation of the Armenian hypothesis is also in order, as it may point to an Iranian origin after all.

buzno m. n. ‘goat’

Sampson: *buznō*. From Iran. *buz* + *-nō*, *-nī*. Cp. Pers. *buz* ‘he-goat’, Afgh.⁹ and Kāf. *vuz*, Kurd. *bizin*. **Boretzky & Igla**: < pers. *buz*. **Vekerdī**: –. **Mānušs**: < viduspers. *buz* ‘goat’. **Calvet**: person, بز (*boz*), s., ‘chèvre’ + *-nō*. **Mégret**: < pers. *boz* ‘bouc’.

This is one of the few uncontroversial Iranian etyma in the lexicon, < Pers. بز *boz*, especially given the presence of etymological [z]. Hancock (1995) also identifies an Afghan¹⁰ cognate *vuz*. It is worth noting that the initial loanword must have been **buz*, from which *buzni* was derived using *-ni*, a specialized suffix for feminine animates, and then regressive oikoclitic derivation occurred *buzni* > *buzno*.

xulaj m. n. ‘master’

Sampson: Fr. Kurd. *xola*. **Boretzky & Igla**: *hulaj*, *khulaj* < pers. *xodā* ‘Herr’, kurd. *xola* ‘Herr, Gott?’. **Vekerdī**: *hulaj*, *xujlo* < kurd. *xola*?. **Mānušs**: < viduspers. *xwatāj* ‘master’, cf. arī. *kadag-xwadāj* ‘householder, master of the house’. **Calvet**: –. **Mégret**: *xulay* < pers.

This word is attested both in Romani (mainly in the Balkan dialect group) and Lomavren. Voskanian (2002) pleads against a Kurdish origin and argues that the transformation [d] > [l] is very common in Lomavren and is a shared isogloss with Romani, but it seems to appear only in some Southern Kurdish dialects. Voskanian also points out that a derivation from Kurdish would be unlikely on semantic grounds, since

⁹ Dari.

¹⁰ Dari.

in Modern Persian خدا *xodā* means ‘God’, losing its primary meaning ‘master’, which is still preserved in both Romani and Lomavren. However, we found that the Kurmanji dialect retains the original meaning: *xudan* ‘owner, master’ (Chyet 2002); *xudî/xūdî* ‘maître, propriétaire’ (Jaba 1879). To wit, we still consider that phonologically the most pertinent explanation is from Pers. خدا *xodā*.

pendex f. n. ‘hazelnut’

Sampson: –. **Boretzky & Igla:** *penex, pexend, pendi*, < kurd. *bendak*; pers. *banduḡ*. **Vekerdî:** < pers. *banduk*. **Mānušs:** *pehend*, čig. dial. *pendex*. < viduspers. *pondik* ‘hazel-nut’. **Calvet:** persan فندق (*bondoq*), فندق (*fandoq*), s., ‘noisette’. **Mégret:** *pehend*, var. dial. *pendex, pendexa* < pers. *pendex*.

The Persian etymon appears to be فندق *fandoq* rather. This is yet another clear-cut instance of the allophonic overlap of [q] and [x] and the adoption of the latter in Romani. Thus, the phonetic transformations taking place would be [x] < [q] and [p] < [f] through fortition of the voiceless labiodental fricative. A possible Turkish filiation (< *findık*) is out of the question, as a rule Pers. [f] > Tur. [f], Pers. [q] > Tur. [k] (Stachowski K. 2020). While the fortition [p] < [f] is plausible, the fricativization [k] > [x] and fronting [ɪ] > [e] would be difficult to account for.

taxtaj m. n. ‘goblet, cup’

Sampson: –. **Boretzky & Igla:** < pers. *tašt, taštak*. **Vekerdî:** < pers. *tašt*. **Mānušs:** < viduspers. *tašt* ‘goblet, cup’. **Calvet:** –. **Mégret:** –.

A Persian etymology is certainly viable < Pers. تشت *tašt* or طشت *tašt* ‘cup, bowl, large vessel’ (Steingass 1892), though difficult to disambiguate from other candidates. Tietze & Lazard (1967) mention an Armenian cognate (*dašd*), as well as a series of Turkish forms (*teşt, teşti, teç, tec*). Hancock (1995) identifies a Kurdish cognate (*test* ‘vessel, trough’). Boretzky (2012) points out that *taxtaj* competes with the native word *khoro* ‘pitcher’, which also designates larger vessels. Notwithstanding, one should not overlook the particular significance that it holds in traditional Roma culture, as a silver goblet or collection thereof (*rupuno taxtaj*) indicates a certain status and respect commanded within the community.

vaxt? m. n. 'time'

Sampson: –. **Boretzky & Igla:** *vakti* [*vahci/ vakci*] Zeit [< alb. *vakt* < tk. *vakit*]. **Vekardi:** –. **Mānušs:** –. **Calvet:** –. **Mégret:** –. **Tahirović:** *vakt* (Eur.), *vakto*. **Lee:** *vakt, vakto* (Eur.). **Courthiade:** *vaxt*. **Sarāu:** *vaxt*.

This entry rather serves as illustration for the lexicographic pitfalls one may succumb to. If *vaxt* were indeed attested among Romani speakers, the presence of the velar fricative [x] would be a very strong proof of a Romani doublet *vaxt, baxt* mirroring almost too fittingly the Persian pair *وقت vaqt* [*vaxt*] 'time' - *بخت baxt* [*baxt*] 'luck' and its allophonic interplay described above. However, this is not the case. *Vaxt* is only recorded as such in the dictionaries by Sarāu (2006) and Courthiade (2009) and it does not seem to reflect existing phonetic realizations.

ROMLEX records the following dialectal variants for 'time': Gurbeti (Banat) – *vrama* (f. n.); Dolenjski – *cajto* (m.), *čas* (m. n.); Finnish Romani – *tijja* (f. n.); Gurvari – *vrama/ vrema/ vreme* (m./f.); kalderaš – *vrama* (f. n.); Latvian Romani – *ciròs* (m.); Lovari – *vrama* (f. n.); Džambazi (Macedonia) – *vakt/ vakti* (m. n.), *vrama* (f. n.); Sepečides – *vakti* (s.m.); Erli (Sofia) – *vakti* (m. n.); Ursari – *timpos* (m. n.); Burgudži – *vahci/ vakci* (m. n.), *vreme* (m. n.); Crimean Romani – *vaxîti* (s. m.); East Slovak Romani – *časos* (m. n.); Gurbet – *vakt/ vakti* (m. n.), *vrama/ vrema/ vrijama* (f. n.); Arli (Kosovo) – *vakti* (m. n.); Lithuanian Romani – *vremja* (f. n.); Arli (Macedonia) – *vahti* (m. n.), *vreme* (m. n.), *saati* (m. n.), North Russian Romani: *vremjo* (m. n.), Romungro (Carpathians): *čas* (m. n.), Sinte – *čiro* (m. n.); Gurbet (Sremski) – *vrama* (f. n.); Veršend – *idejo* (m. n.). Note also the form *vakici* in the Spoitori dialect (Sarāu 1998).

We can note the presence of two main isoglosses, a Turkish one in the Balkan dialects (*vakt, vakti, vakici, vahti, vaxîti*) and a Slavic one in the 'Continental' dialects (*vrama, vrema, vrijama, vreme, vremjo, vrama*), with a middle overlap area. Hancock (1995) records the Balkan isogloss *vaxt* and *vakti*, and observes that even though this word appears in Persian (*وقت vaqt*), "its form suggests Turkish as immediate source of adoption in both languages". However, it appears that in Turkish *vakit* is an early Arabic loan (as well as its synonym *zaman*), and so is its Persian counterpart. Iranian linguists do not consider it to be a Turkish loanword (Ershadi 2014). In Turkish, phonological nativization occurred through the velarization of the uvular stop [q] > [k].

zet m. n. 'oil'

Sampson: –. **Boretzky & Igla:** *zejtin* [zejti, zetin] [< tk. bzw. arab-pers. *zejtīn*]. **Vekardi:** –. **Mānušs:** –. **Calvet:** *zetino* S.-c.¹¹, *zejtin*, s.m. **Mégret:** –. **Tahirović:** –**Lee:** *zétino*. **Courthiade:** *zet*.

This is a typical Turkish-(inter)mediated Farsism and it only appears in some dialects from the Balkan group. Hancock (1995) is the only one including it in his glossary of Iranisms, unjustifiably so.

6. Conclusions

In view of the sample cases treated above, we may draw some preliminary conclusions regarding the Iranian component in the Romani lexicon.

Phonologically, the voiceless velar fricative [x] is indeed a useful diachronic marker of post-Indian loans, and in particular cases it points to a Persian origin. The presence of etymological [x] in words such as *baxt* and *xulaj* adds weight to their Persian filiation. Moreover, when in auslaut, [x] seems to derive occur almost exclusively from Persian, as shown above for *bezex*, *buzex*, and *pendex*. In the same realm of historical phonology, the findings of eminent Turcologists such as Pomorska (2013) and Stachowski K. (2020) provide useful guidance for confirming or ruling out Turkish and/or Balkan L2 mediation.

Cross-dialectal lexicographic research is essential for mapping out the distribution and stability of certain lexemes and isoglosses. This helped us weed out prescriptivist lexicographic entries such as *vaxt* (Sarău 2006, Courthiade 2009) or Turkish/L2-mediated words with limited dialectal distribution such as *zet*.

We anticipated that a contrastive comparison with the Lomavren corpus would be of limited applicability, which does not, however, detract from its usefulness. In the case of *bezex*, it provides further arguments for a convincing refutation of a possible Turkish mediation. The regular [d] > [l] sound change, which is a shared isogloss with Romani, might reinforce the likelihood a Persian etymology of *xulaj*, even though a Kurdish origin cannot be discounted outright.

¹¹ Serbo-Croatian.

In summary, a more comprehensive and multi-faceted lexicographic and etymological approach can alleviate somewhat the challenges one is faced with when trying to discern and disambiguate the Iranian lexicon in Romani. Future research should certainly include a more expansive approach and further refine its focus. A more robust corpus should be explored and research should branch out into other Iranian languages or dialects that may have had a lexical impact on Romani.

BIBLIOGRAPHY

- Achim, Viorel (2004), *The Roma in Romanian History*, Central European University Press, Budapest, New York.
- Apte, Vaman Shivram (1993), *The Student's Sanskrit-English Dictionary*, Motilal Banarsidass, Delhi.
- Aryanpur Kashani, Manoochehr (2017), *The Aryanpur Progressive English-Persian Persian-English Dictionary (Comprehensive)*, Computer World, Teheran.
- Baló, Márton A. (2020), "Romani Phonology", in Matras Y. & Tenser A. (eds.), *The Palgrave Handbook of Romani Language and Linguistics*, Palgrave Macmillan, pp. 119-153.
- Boretzky, Norbert, Igla, Birgit (1994) *Wörterbuch Romani-Deutsch-Englisch für den südosteuropäischen Raum: mit einer Grammatik der Dialektvarianten*, Harrassowitz, Wiesbaden.
- Boretzky, Norbert (2012), *Studien zum Wortschatz des Romani*, Faber, Veliko Tarnovo.
- Calvet, Georges (2021), *Dictionnaire tsigane-français: dialecte kalderash*, L'Asiathèque – Maison des langues du monde, Paris.
- Chyet, Michael L. (2002), *Kurdish-English Dictionary. Ferhenga Kurmancî-Inglîzî*, Yale University Press, New Haven and London.
- Clauson, Gerard (1972), *An Etymological Dictionary of Pre-Thirteenth-Century Turkish*, Clarendon Press, Oxford.
- Courthiade, Marcel (2009), *Morri Angluni Rromane Ćhibăqi Evroputni Lavustik*, Romano Kher, Budapest.
- Courthiade, Marcel (2016), "Knowledge Based on Sources and Historical Data versus Knowledge Based on Clichés and Legends in the Indian Stage of Rromani History", in *Analele Universității "Ovidius" din Constanța. Seria Filologie*, XXVII(2), pp. 237-282.
- Ershadi, F.A. (2014), *Farhang-e vājgan-e tarki dar zabān va adabiyāt-e fārsi. Turkish Lexicon in Persian [Language and] Literature*, Entesharat-e Bagh-e Andishe, Erdebil.
- Finck, Franz Nikolaus (1905), "Die Sprache der Armenischen Zigeuner", in *Mémoires de l'Académie Impériale des Sciences de St.-Pétersbourg*, VIII^e Série, Volume VIII, N. 5, St. Pétersbourg.

- Friedman, Victor A. (1989), "Toward defining the Position of Turkisms in Romani", in Šipka M. (ed.), *Jezik i Kultura Roma*, Institut za proučavanje nacionalnih odnosa, Separat, Sarajevo, pp. 251-267.
- Friedman, Victor A. (2020), "The Impact of Turkish on Romani", in Matras Y. & Tenser A. (eds.), *The Palgrave Handbook of Romani Language and Linguistics*, Palgrave Macmillan, pp. 119-153.
- Ghavami, Golnaz Modarresi (2018), "Phonectis", in Sedighi A. & Shabani-Jadidi P. (eds.), *The Oxford Handbook of Persian Linguistics*, Oxford University Press, Oxford, pp. 128-151.
- Hancock, Ian (1995), "On the migration and affiliation of the Dōmba: Iranian words in Rom, Lom and Dom Gypsy", in Matras Y. (ed.), *Romani in contact. The history, structure and sociology of a language*, John Benjamins, Amsterdam/Philadelphia, pp. 25-51.
- Jaba, Auguste (1879), *Dictionnaire Kurde-Français*, Imprimerie de l'Académie Impériale des Sciences, St-Petersburg.
- Johanson, Lars (2010), "Turkic Language Contact", in Hickey R. (ed.), *The Handbook of Language Contact*, Wiley-Blackwell, Malden, MA, pp. 653-672.
- Lee, Ronald (2010), *Romano Alavari Kalderashitska-Inglezitska. Romani Dictionary Kalderash-English*, Magoria Books, Toronto.
- Maggi, Mauro, Orsatti, Paula (2018), "From old to new Persian", in Sedighi A. & Shabani-Jadidi P. (eds.), *The Oxford Handbook of Persian Linguistics*, Oxford University Press, Oxford, pp. 26-78.
- Mānušs, Leksa, Neilands, Jānis, Rudevičs, Kārlis (1997), *Čigānu-latviešu-angļu etimoloģiskā vārdnīca un latviešu-čigānu vārdnīca. Romany-Latvian-English etymological dictionary and Latvian-Romany dictionary*, Zvaigzne ABC, Riga.
- Matras, Yaron (2004), *Romani. A Linguistic Introduction*, Cambridge University Press, Cambridge.
- Mégret, Jean-Claude (2016), *Dictionnaire de la Romani Commune (langue tsigane)*, L'Harmattan, Paris.
- Parker, Philip M. (2008), *Webster's Turkish-English Thesaurus Dictionary*, ICON Classic, San Diego.
- Pisowicz, Andrzej (1985), *Origins of the New and Middle Persian Phonological Systems*, Uniwersytet Jagiellński, Krakow.
- Pomorska, Marzanna (2013), *Materials for a Historical Dictionary of New Persian Loanwords in Old Anatolian and Ottoman Turkish from the 13th to the 16th Century*, Jagiellonian University Press, Kraków.
- ROMLEX, Online comparative lexical database of Romani, <http://romani.uni-graz.at/romlex>.
- Sampson, John (1926), *The Dialect of the Gypsies of Wales being the older form of British Romani preserved in the speech of the clan of Abram Wood*, Clarendon Press, Oxford.
- Sarău, Gheorghe (1998), *Dicționar rrom (spoitoresc)-român*, Kriterion, București.
- Sarău, Gheorghe (2006), *Dicționar rrom-român*, Editura Sigma, București.
- Sarău, Gheroghe (2012), *Dicționar român-rrom*, Editura Sigma, București.
- Scala, Andrea (2014), "The mixed language of the Armenian Boshia (Lomavren) and its inflectional morphology: some considerations in light of Armenian dialectal variation", in *Annali del Dipartimento di Studi Letterari, Linguistici e Comparati Sezione linguistica*, Università degli studi di Napoli "L'Orientale", N.S. 3, pp. 233-250.

- Stachowski, Kamil (2020), "Phonetic Renderings in Turkish Arabisms and Farsisms", in *Türkkbilig*, 2020/40, pp. 23-47.
- Stachowski, Marek (2019), "Slavic languages in contact, 2: Are there Ottoman Turkish loanwords in the Balkan Slavic languages?", in *Studia Linguistica Universitatis Iagellonicae Cracoviensis*, 136, pp. 99-105.
- Steingass, Francis Joseph (1892), *Persian-English Dictionary. Including the Arabic words and phrases to be met with in Persian Literature*, Librairie du Liban Publishers, Beirut.
- Tahirović-Sijerčić, Hedina (2013), *Romano Alavari Gurbetcko-Englecko/ Englecko-Gurbetcko. Romani Dictionary English-Gurbeti/ Gurbeti-English*, Magoria Books, Toronto.
- Tietze, Andreas, Lazard, Gilbert (1967), "Persian Loanwords in Anatolian Turkish", in *Oriens*, Vol. 20, pp. 125-168.
- Turner, Ralph Lilley (1926), "The position of Romani in Indo-Aryan", in *Journal of the Gypsy Lore Society*, Third series, V (4), pp. 145-189.
- Vekerdi, Jozsef (2000), *A magyarországi cigány nyelvjárások szótára. A Comparative Dictionary of Gypsy Dialects in Hungary*, Terebess Kiadó, Budapest.
- Voskanian, Vardan (2002), "The Iranian loanwords in Lomavren, the secret language of the Armenian Gypsies", in *Iran & the Caucasus*, 6 (1-2), Brill, Leiden, pp. 169-180.
- Windfuhr, Gernot (2009), "Dialectology and Topics", in Windfuhr G. (ed.), *The Iranian Languages*, Routledge, London, pp. 5-42.
- www.tureng.com