



Universiteit
Leiden
The Netherlands

Some plant and animal names in Gavruni

Bernard, C.B.A.S.; Garnier, R.

Citation

Bernard, C. B. A. S. (2020). Some plant and animal names in Gavruni. In R. Garnier (Ed.), *Innsbrucker Beiträge zur Sprachwissenschaft* (pp. 27-61). Innsbruck: Innsbrucker Beiträge zur Sprachwissenschaft. Retrieved from <https://hdl.handle.net/1887/3220779>

Version: Publisher's Version

License: [Licensed under Article 25fa Copyright Act/Law \(Amendment Taverne\)](#)

Downloaded from: <https://hdl.handle.net/1887/3220779>

Note: To cite this publication please use the final published version (if applicable).

TABLE OF CONTENTS

1 - Milad ABEDI, University of Zurich “On the later phase of Elamite-Iranian language contact”	1
2 - Chams BERNARD, Leiden University “Some plant and animal names in Gavruni”	27
3 - Michele BIANCONI, University of Oxford “Some thoughts on Anatolian lexicon in Mycenaean Greek”	63
4 - Hugo BLANCHET, Université de Limoges, ÉPHÉ “Méfitis osque et Méfitis romaine, des sources limpides aux eaux pestilentiellles”	89
5 - Václav BLAŽEK, Masaryk University, Brno “Latin <i>bellua/bēlua</i> ‘beast’ of Celtic origin?”	113
6 - Gerd CARLING, Lund University, “A dangerous story: the linguistic behaviour of the category sharp cutting implements”	121
7 - Xavier DELAMARRE, CNRS – AOROC “Les noms gaulois dans l’onomastique impériale”	151
8 - Romain GARNIER, Université de Limoges, IUF & Benoît SAGOT, INRIA “New results on a centum substratum in Greek: the Lydian connection”	169
9 - Jean HADAS-LEBEL, Université Lumière Lyon 2 “Une origine étrusque pour lat. <i>corōna</i> ?”	201
10 - Daniel KÖLLIGAN, Julius-Maximilians-Universität Würzburg “Deaffrication in Armenian?”	215
11. Martin KÜMMEL, Friedrich-Schiller Universität Jena “Substrata of Indo-Iranic and related questions”	237

VI

Table of contents

12 - Jean-Pierre LEVET, Université de Limoges “Des traces d’un ancêtre eurasiatique en indo-européen ?”	279
13 - Rosemarie LÜHR, Friedrich-Schiller-Universität Jena Honorarprofessorin an der Humboldt-Universität zu Berlin “Zum Langobardischen als Trümmersprache”	295
14 - Vincent MARTZLOFF, Sorbonne Université “La question du <i>substrat sicule</i> dans le Latium. Souvenir authentique ou mythe historiographique”	315
15 - Ranko MATASOVIĆ, University of Zagreb “Language of the bird names and the pre-Indo-European substratum”	331
16 - Veronika MILANOVA, University of Vienna Sampsa HOLOPAINEN, University of Helsinki Jeremy BRADLEY, University of Vienna “Contact phenomena in IE kinship and social terms and beyond”	345
17 - Georges-Jean PINAULT, ÉPHÉ, PSL “Tocharian lexicon in the light of contact phenomena”	367
18 - Benoît SAGOT, INRIA “A new PIE root <i>*h₁er-</i> ‘(to be/become) dark red’”	403
19 - Zsolt SIMON, Ludwig-Maximilians-Universität München “Die hurritischen Lehnwörter im Keilschriftluwischen”	411
20 - Aljoša ŠORGO, Leiden University “Characteristics of lexemes of a substratum origin in Proto-Germanic”	427
21 - Dan UNGUREANU, Charles University, Prague “The four layers of the lexical substrate in Romanian”	473
22 - Michael WEISS, Cornell University “The plough and its parts in western Europe”	481
INDEX	501

FOREWORD

The conference on “Loanwords and Substrata” took place on June 4th–7th, 2018 in the Faculty of Letters and Humanities at the University of Limoges. 30 papers were read: 22 of them have been submitted for these Proceedings. In the present volume, 15 papers are written in English, 5 in French and 2 in German. The conference topic was the question of loanwords – whether they are known as such or not – and the notion of substrata, including semantic or even syntactic calques, in ancient or modern Indo-European languages, with a special emphasis on contacts between Indo-European and Non-Indo-European languages such as Hurrian, Elamite, Etruscan, Ugaritic, Finno-Ugrian and Austroasiatic languages.

The question of the *Völkerwanderung* is addressed in several papers, within the scope of accounting for the substratic part of the lexicon, pointing to ancient contacts of people, engaged in wars or in trade. Some papers are comparable to a short dictionary, while other papers provide a complete survey of any possible contacts between a whole linguistic group and its neighbouring languages. Some studies are dedicated to a single word, offering a totally new reassessment of its etymology. Long-range issues are also addressed, such as the contacts between IE and Finno-Ugrian, or the possibility of a substratic language in the so-called “*Nordwestblock*”, admittedly reflected by agricultural or animal names.

Interestingly, one may compare alternative approaches concerning a single lexeme or a suffix: for instance KÜMMEL’s (258) and BERNARD’s views (37ff.) concerning Ilr. *-ā́ca-, or PINAULT’s (391) and WEISS’ (484) contradictory explanations of Toch. B *āre* ‘plough(ing)’. Lat. *ardea* [f.] ‘heron’ is accounted for as a substratic word by MATASOVIĆ (339) and ŠORGO, (433–434), while SAGOT (407) proposes a PIE origin for this obscure bird name.

Anatolian languages are dealt with: the cultural borrowings from Hurrian in CLuwian are studied by Zsolt SIMON, “Die hurritischen Lehnwörter im Keilschriftluwischen” (411–426). The crucial Graeco-Anatolian connection is tackled by Michele BIANCONI, “Some thoughts on Anatolian lexicon in Mycenaean Greek” (63–88) and by Romain GARNIER & Benoît SAGOT, “New results on a centum substratum in Greek: the Lydian connection” (169–200). Three innovative studies encompass various issues of the huge Indo-Iranian field: Martin KÜMMEL (keynote speaker), “Substrata of Indo-Iranic and related questions” (237–277), a long-range survey of the contacts of “Indo-Iranic” with no fewer than eleven language families, with original considerations on the Mitanni Ilr. material. Milad ABEDI provides a study “On the later phase of Elamite-Iranian language contact” (1–25), while Chams BERNARD makes a full reassessment of “Some plant and animal names in Gavrūni” (27–61), a long inquiry involving many languages, such as Proto-Malayo-Polynesian or Mon-Khmer.

The Italic field is well represented by Michael WEISS (keynote speaker), “The plough and its parts in western Europe” (481–500). Václav BLAŽEK gives a new etymological proposal in “Latin *bellua/bēlua* ‘beast’ of Celtic origin?” (113–119). Vincent MARTZLOFF deals with “La question du *substrat sicule* dans le Latium. Souvenir authentique ou mythe historiographique” (315–330). Hugo BLANCHET makes a study on “Méfitis osque et Méfitis romaine, des sources limpides aux eaux pestilentielles” (89–112). Jean HADAS-LEBEL provides a new explanation for a very famous word: “Une origine étrusque pour lat. *corōna* ?” (201–213), while Dan UNGUREANU makes a thorough survey of “The four layers of the lexical substrate in Romanian” (473–480). Germanic is not left behind, with Rosemarie LÜHR’s paper “Zum Langobardischen als Trümmersprache” (295–314) and Aljoša ŠORGO’s massive study on the “Characteristics of lexemes of a substratum origin in Proto-Germanic” (427–472). Daniel KÖLLIGAN makes a reassessment on “Deaffrication in Armenian” (215–235), Xavier DELAMARRE deals with “Les noms gaulois dans l’onomastique impériale” (151–168). Georges-Jean PINAULT makes new considerations on “Tocharian lexicon in the light of contact phenomena” (367–401). Some studies involve the PIE lexicon itself: Gerd CARLING, “A dangerous story: the linguistic behaviour of the category sharp cutting implements” (121–149) and Benoît SAGOT, “A new PIE root **h₁er-* ‘(to be/become) dark red’” (403–409). Some papers address issues concerning PIE and beyond, such as Jean-Pierre LEVET’s “Des traces d’un ancêtre eurasiatique en indo-européen ?” (279–294) or Ranko MATASOVIĆ (keynote speaker), “Language of the bird names and the pre-Indo-European substratum” (331–344) and Veronika MILANOVA, Sampsa HOLOPAINEN & Jeremy BRADLEY, “Contact phenomena in IE kinship and social terms and beyond” (345–366).

ACKNOWLEDGEMENTS

The printing of this book was supported by the Institut Universitaire de France (IUF) and the gathering was supported by the University of Limoges, with the assistance of Chams BERNARD and Benoît SAGOT, whose welcome help could not be sufficiently commended. I also have to express my deepest gratitude to Chams BERNARD and Xavier DELAMARRE, who helped me greatly for the index, and to my students, who have been more than helpful. Last but not least, I thank for their willingness Corinne SYLVESTRE and Céline CHRÉTIEN, kingpins of the administrative centre of Limoges University.

Romain GARNIER



LW 121, in Tarris kürzlich gefunden

Chams Benoît BERNARD
Leiden University

Some plant and animal names in Gavruni

ABSTRACT. — This paper presents nine Gavruni animal and plant names, and their various cognates within and without Iranian. It is etymological in purpose, and proposes new etymologies for some words.

1. Introduction¹

Various lexical studies have demonstrated that lesser studied Iranian languages conceal great riches.² These studies have mostly concentrated on archaisms, and how archaisms found in modern Iranian languages can help us improve our understanding of the different stages of Iranian languages, in particular Proto- and Old Iranian. In the present study, I will look at archaisms in the inherited lexicon as well as borrowings and substratum words. In order to do so, I have selected a number of Gavruni plant and animal names; I will discuss their etymology and, when needed, their meaning. In this category of the lexicon, effects of contact and archaic loans can be seen, but one also finds inherited vocabulary. Some of these studied words, having no straightforward etymology and seemingly old, are probably of substratal origin, some are of Indo-European origin, while others are onomatopoeic. Three of the words studied here (*hros*, *kalpak*, *konjet*), have apparent correspondents in Sanskrit or Old Iranian, and so might belong to the Indo-Iranian substratum posited by LUBOTSKY (2001). This substratum has mostly been studied from the point of view of Indian languages, in particular Sanskrit (WITZEL 1999, LUBOTSKY 2001), but I believe much information can be yielded from looking at the Iranian side. Most of the etymologies in the present study are tentative, based on my previous study of Gavruni historical phonology (BERNARD 2016), and various other works on Gavruni and Iranian languages. Gavruni (Gav.) is a modern Iranian language. It is the traditional language of the Zoroastrians of Iran, and was spoken by all Iranian

¹ This research was supported by the Netherlands Organisation for Scientific Research (NWO, project number 276-70-028). I am very grateful to Dr. Kate BELLAMY for thoroughly correcting the English. My heartfelt thanks go to Prof. Gilles BERNARD, Dr. Alessandro DEL TOMBA, Dr. Guillaume JACQUES, Prof. Martin KÜMMEL, Prof. Alexander LUBOTSKY, Dr. Michaël PEYROT and Prof. Martin SCHWARTZ for their suggestions, corrections and valuable comments. I am grateful to Dr. Yvonne VAN AMERONGEN for spending time discussing botany with me. I also thank Dr. Romain GARNIER for editing this volume with so much care and attention to detail. All remaining mistakes and misconceptions are my own.

² For instance, VOSKANIAN & BOYAJIAN-SURENIANTS (2007); ASATRIAN (2009).

Zoroastrians (in the regions and cities where Zoroastrians remained) until the middle of the 20th century. Since this time it gradually became extinct in the Zoroastrian community of Kermān, but remained steady elsewhere. In 2016 there were only three competent speakers of the Kermāni dialect (GHOLAMI & FARAHMAND 2016), of whom only one remains today. Gavruni is nowadays mostly spoken in Yazd, a city in the desert, and, to a lesser extent, in Tehran and Shiraz, in total by around 10,000 speakers.

Gavruni speakers in the past (that is, until the late 1970s) were mostly farmers, but a number of them were merchants or worked in various trades. One major occupation for female Gavruni speakers was weaving, but they also took part in farming activities, at least occasionally, and in one village, even as much as men (on this last point, cf. BOYCE 1969:126). Gavruni speakers in and around Yazd are traditionally spread out between Yazd and its suburbs, some smaller cities and a number of small Zoroastrian villages.

There are at least nineteen recognizable varieties of Gavruni spoken today (BERNARD 2016:16), plus Kermāni, and a number of dead varieties from the now deserted villages around Yazd. Unfortunately, only Malati (the most prestigious dialect of Yazd), and the Kermāni dialect have been studied extensively to date. My 2016 study (BERNARD 2016) was the first to systematically take dialectal data into account for historical purposes. For the present study, I will mostly use lexicographic sources since they yield more archaic forms, with more accurate meanings, than many later works.

I will base myself mainly on SORUŠYĀN 1978 (henceforth SFB), a dictionary of Gavruni written in Persian. It gives Yazdi forms (always from the Malati dialect), Kermāni forms, and more rarely some Tehrani forms, as well as some “village forms” from around Yazd. It is written in Persian, so the Gavruni words are written in the Arabo-Persian alphabet, and consistently transcribed in the Latin alphabet. It contains a great number of agricultural words, plant and animal names, along with a treasure trove of almost forgotten archaic Gavruni words. I will secondarily use other sources, such as HOUTUM-SCHINDLER 1882, an article on the Zoroastrians of Iran that contains a number of Kermāni Gavruni words translated into German, and BOYCE 1969, an article on farming in Šarif-ābād (Gav. *Šarpavo*), a Gavruni-speaking village in the surroundings of Ardakān (a city north to Yazd), which contains a list of agriculture-related words in that dialect. Finally, for two words, I use GHOLAMI & FARAHMAND 2016, a manual of Kermāni Gavruni, and for one, MAZDĀPOUR 2006, a dictionary of the Malati dialect of Gavruni. If all or almost all sources agree on a particular form, I indicate “various sources” or “all sources” next to the Gavruni word. I

cite the Kermāni Persian words from SOTOUDEH 1956. This study will naturally lead to the consideration of a number of non-Gavruni forms, Iranian or not. My ultimate goal in presenting these words is to show that the variety of strata of vocabulary in Modern Iranian languages is of great interest and goes beyond the usual inherited, Arabic, Turkic or Persian set of etymologies.

2. *poxtu* Kermān Gav. ‘فاخته’ = ‘turtledove’ (SFB)

Originally a borrowing from Arabic *fāxta* ‘turtledove, cuckoo’, or from Classical Persian *fāxta* ‘id.’, itself a borrowing from this Arabic word. Classical Arabic has *fāxīta* but in most Arabic dialects, *ī* is dropped in open syllables. Word initial *f* is borrowed as *p* in Gavruni in older (prehistorical) borrowings, long *ā*, realized /ɒ/, is secondarily rounded to Proto-Gavruni *ɔ, which became /o/ in some positions in Kermāni Gavruni (BERNARD 2016:78–79).

SCHAPKA (1972:180) proposed that Arabic *fāxīta* derives from an Iranian language.³ “فاختک *fāxtak* ‘Ringeltaube’ *Columba palumbus*, ‘Holztaube’ [...] Arab. *fāxīta* ‘Taube’ ist angeblich Lehnwort aus dem Persischen, und *fāxta* (mit ostiran. β- (ف) < b-) stammt möglicherweise von der W[ur]z[el] idg. *bhag- (POKORNY p. 107) in einer ihrer vielen Bedeutungen ab.” This etymology is difficult for multiple reasons: phonetically, a Middle Iranian form *βaxtag, *βaxtay or *βaxtak would yield Arabic †waxtag/q/k or perhaps †fāxtag/q/k (cf. EILERS 1971:601), but not *fāxīta*. Semantically, I am not convinced that any of the meanings shown in POKORNY: ‘zuteilen’, ‘scharf, auch von Geschmack’, ‘Buche’ represent plausible etymologies for ‘turtledove, cuckoo’. These arguments make the hypothesis of an Iranian origin of this word doubtful.

In fact, it seems much more plausible that the Arabic word derives from Arabic *faxt* (الْفَخْتُ) ‘moonlight’, designating the bird “because of its colour” (see LANE 1865:2348).⁴ There is yet another word designating the same bird in another Arabic dialect, قُمْرِيّ /qumriyy/, “so designated because أَقْمَر in colour” (it designates “a species of collared turtle-dove, of a dull white colour marked with a black collar [...]”) LANE (1865: 2563). One of the meanings of أَقْمَر *aqmar* is “[...] a dull or dingy or dusky white” (id.), and the word is evidently related to the root *q-m-r* ‘moon’. It was borrowed, for example, in Khwarezmian <qmry>, <qmryk> ‘Turteltaube’ (BENZING 1983:549). These two similar etymologies confirm, in my view, the Arabic origin of Persian *fāxta* and Gav. *poxtu* < **fāxta* + *-ū.

³ So EILERS (1971:618).

⁴ There are also two Arabic verbs *faxata* ‘to walk elegantly’, and *faxata* ‘to coo’, said of the (turtle) dove.

The *-u* in *poxtu* corresponds to the *-u* nominal suffix of Kermān Gavruni (historically from OIr. **-ū-ka-*), cf. *toxm* ‘egg, seed’ (a borrowing from NP *toxm*, cp. inherited Gavruni form *tum*) → Kerm. *toxmu* ‘watermelon seed used at the end of a spindle’ (SFB), Kerm. *teterk* ‘hail’ → *teterku* ‘id.’.⁵ It is unclear to me whether the form *poxtu* is derived from Proto-Gavruni (PG) **poxta* or from **poxtak* < *fāxta* with an *-ak* suffix. In both cases, there would be a secondary (re-)suffixation. See *totorog* (p. 55) for the current Yazdi equivalent.⁶

3. *gok, guk* Kermān Gav. ‘Frosch’, قورباغه, وزغ = ‘frog, toad’ (HOUTUM-SCHINDLER, SFB), *vak* Yazd Gav. ‘id.’ (SFB)

These words should be examined separately. The first one, *gok* ~ *guk*, is quite similar to the Persian name of the toad: غوک /*ɣo:k*/ or /*ɣu:k*/, Tajiki Ҷук (*ɣuk*) (BAIZOYEV & HAYWARD 2004:203; 332) according to HENNING (1939:95) a borrowing from Sogdian *ɣwk*. Whether Gav. *gok* ~ *guk* is borrowed from Persian, or from another language, seems to be an unsolvable problem.

Yazd Gav. *vak*, on the other hand, corresponds to Middle Persian *wak*, New Persian *bak*, Baxtiari *baq*, South Baškardi *vak*, North Baškardi *gwak*, all these forms stemming from **wak* (for the Baškardi data, see VOSKANIAN & BOYAJIAN-SURENIANTS 2007:124).

It is possible that the Iranian lexeme was originally **wak*, and became **gwak* in one of the Iranian languages that underwent *#w- > #gw-*, such as Balochi and North Baškardi, and that this **gwak* spread to numerous languages, becoming *gok, guk* in Persian (Sogdian *ɣwk*) could also represent /*ɣwak*/, and so could Pers. غوک). Kerm. Gav. *gok, guk* could be a borrowing from Persian /*ɣo:k*/ and /*ɣuk*/ (the former being older than the latter), respectively, and Yazd Gav. *vak* could represent an “inherited” form.⁷ Of course, and perhaps more likely, *gok*

⁵ Gav. *teterku* also designates a type of harmless pox (probably chicken pox). I do not know whether there is a semantic link between ‘hail’ and ‘chicken pox’: maybe the spots of pox resemble those left by a rain of hail on someone? Hail is naturally not a very common phenomenon in Kermān.

⁶ PETERMANN wrote the following about Yazd and its Gavruni-speaking population (1865:207): “Es giebt auch hier, wie in dem ganzen Orient, viele Lauchtauben mit grau-braunem Gefieder und einem schwarzen Ring um den Hals, die man puchta nennt; in Hille nannte man sie fuchte, aber das arabische f scheint man hier gar nicht zu kennen, daher man auch pursi für fāresi spricht.” This shows that there was an old Yazd Gavruni word *poxta* (transcribed as *puchta*) by PETERMANN) designating that particular species of dove.

⁷ While both forms are likely to be onomatopoeic in origin (so VOSKANIAN & BOYAJIAN-SURENIANTS 2007:124), the forms *gok* and *guk* (and *ɣwk*?) might – at least in some Iranian language(s) – have been influenced by **gayaka-* ‘small cow’, referring to one particular frog, the Iranian tree frog, *Hyla savignyi*, and in particular to its voice. “They are the smallest frogs in Iran,

and *gwak* etc. are simply onomatopoeic words, imitating the sound of the frog, cf. the onomatopoeias *gwac'h* ([gʷax]) in Breton and *kvak kvak* in Slovenian.

4. *gok som* Kermān Gav. ‘آبدزدک’ = ‘mole cricket’ (SFB)

While *gok* here certainly means ‘frog’ (cf. Section 3), the meaning of *som* is not immediately evident. I believe it is related to NP *suftan*: *sunb-/suft-* ‘to pierce or dig a hole’ (cf. newly remade weak verb Contemporary Pers. *sonbidan* ‘id.’), and is either the Gavruni name of the ‘mole’ (the ‘digger’?) or, and probably more likely, a post-verbal adjective of that same verb.⁸ The compound would then mean something along the lines of ‘mole frog’, ‘frog digger’, or rather ‘digger-frog’. Mole crickets are digging animals, often considered as agricultural pests (especially the immigrant species). They can bite human beings, but their bite is harmless. They dig into the ground with their two forelegs and have burrows under the ground, hence their name.⁹ They were perhaps qualified as ‘frogs’ because they belonged to the Zoroastrian category of “frogs”, cf. *Greater Bundahišn* XXII:17. This is maybe reminiscent of the compounds with *go* ‘cow’ in Gavruni, see Section 6.4.2. (p. 40).

The Persian name of mole crickets, آبدزدک (*āb-duzdak*), literally means ‘small water-thief’. This is perhaps because they can damage dykes and small streams.

5. *hros* Yazd Gav. (various sources), Kermān Gav. *oros* (GHOLAMI & FARAHMAND 2016:159) and *orus* (HOUTUM-SCHINDLER 1882:63) ‘rooster’, Yazd Gav. (Malati dialect) *xorusak* ‘chick’ (MAZDĀPOUR 2006:431)

5.1. Cognates of the Gavruni forms

There are many cognates of Proto-Gavruni (PG) **hros* ‘rooster’ in other Iranian languages: MP *xrōs* (CPD 94), Cl. Pers. *xurōs*, both borrowed from another Iranian language, and Cl. Pers. *xurōh*, the expected inherited form, Xavānsari

but they have the loudest voices, and whenever a frog chorus fills the night air, one can be sure it is this tree frog.” (ANDERSON 1985). This usage would be reminiscent of the English word *bullfrog*, designating frogs whose voice is strong and deep, like that of bulls and cows. Since this particular species of frog is not common in Kermān, this hypothesis does not directly concern these Gav. forms, even more so considering the fact that cows are sacred in Zoroastrianism, while frogs are despised.

⁸ The sound change *unb* > *um* is attested in Persian, as is the word *sum* < *sunb* (both *sumb* ‘hoof’ > *sum*, and *sumb* ‘digging, digger’ > *sum* > Cont. Persian *som*).

⁹ Another reason for them to be called *mole crickets* (French *taupe-grillon*) is that the fore part of their body looks like that of a mole.

krōs, Balochi *krōs* and *krōsk* (KORN 2005:118; 165), Gilaki *xorus*, Farizandi *xarūs*, Yārāni *harūs*, Natanzi *xorūs* (cf. CHRISTENSEN 1930:287), Sistāni *xross* (XOMAK 2000:170). MORGENSTIERNE (1938:268) connects the Yidgha word *xurúso* (< OIr. **xrausā-*) to these words. The meaning of *xurúso* is quite different: “n[ame] of an animal resembling a fox, but larger, which attacks ibexes, Cyon Alpinus Pallas?”. There is also a Sanskrit word which is often proposed as a cognate of the Iranian words for ‘rooster’: Ved. *kroṣṭár-* m. ‘Schakal’ (EWAia¹ 416). The latter meaning is closer to that of the Yidgha word than to that of ‘rooster’.

All of these words are usually connected to the Indo-Iranian root **krauc-* ‘to call’ (cf. CHEUNG 2007:448–449), Ved. *kroś-*, Av. *xraos-*, Khot. *grūs* ‘to call’ (DKS 93), MP *xrōstan* (CPD 94) ‘to call, cry’, Cl. Pers. *xurōšīdan* ‘id.’, etc.

Ved. *kroṣṭár-* is a *-tar-* derivative meaning ‘the crier, the caller’, and it will be quite obvious to the reader why roosters are called ‘criers’ or ‘callers’. The Sistāni form *xross* derives from Pre-Sistāni **xrust*. The reconstructed form **xrust* can hardly go back to a *-tar-* derivative, since a full grade would be expected for such a form (†*xrōss*).¹⁰ Sistāni *xrō* (XOMAK 2000:170) is a borrowing from Cl. Pers. *xurōh*. Finally, perhaps Sanskrit *kruñc-*, *krauñca-*, m. ‘crane’ is also related, see Section 5.3. Cranes too make loud, characteristic sounds. The previous etymologies of *kruñc-*, *krauñca-* have been discussed in HAMMER (2019:401–403), who himself proposes that these forms are onomatopoeic in origin, and thus does not account for the form in detail.

5.2. Discussion of the Gavruni forms

Both *hros* and *oros*, the Yazd and Kermān Gavruni forms, derive from a form starting with a cluster **hr-*. I would rather reconstruct the Proto-Gavruni form as **hros* and take the form *orus* as influenced either by Contemporary Persian *xorus* or Classical Persian *xurōs*, due to the vowel /u/ in *orus* (< PG **ū* or **ō*), an influence typical of Kermān Gav. To support this reconstruction, one should consider that the form *oros* given much more recently by GHOLAMI & FARAHMAND cannot be explained by contact phenomena, and thus needs to be more archaic than *orus*.

The Kerm. forms *oros* and *orus* show that the cluster epenthesis preceded the loss of word-initial *h-* in Kermān Gavruni: PG **hros* > Pre-Kerm. **horos* > Kerm. *oros*.

¹⁰ One could also speculate that the Ved. form *kroṣṭár* also goes back to a preform nom. **krućtá-* which was reinterpreted as a *-tar-* noun, hence also the full grade. There is nevertheless no evidence at all for this.

Finally, Yazd Gav. (Malati dialect) *xorusak* ‘chick’ is borrowed from Cont. Persian *xorusak* ‘idem’.

5.3 Further etymology

If all the words cited in 5.1 refer to animals with loud voices, callers and criers, and since not a single animal name can be reconstructed for the Indo-Iranian proto-form of all these cognates, it is probable that the meaning of the reconstructed root was still ‘the crier, the caller’. The further etymology of the root and its derivation deserves particular attention. It has been suggested by some (as mentioned in, for example, EWAia¹ 416) that PIIr. **krauć-* has Indo-European cognates besides Indo-Iranian, mainly Lithuanian *kraũkti* ‘to croak’, *krauklỹs* ‘crow’. This nevertheless does not seem possible, since **-k* does not alternate with **-ĥ* in Proto-Indo-European. The Cl. Pers. form *xurōšīdan* ‘to call, cry’, possibly denominative of Cl. Pers. *xurōš* ‘call, cry’ (vs. Parthian *xrōs* ‘idem’),¹¹ might indicate that this root was initially **krau-* enlarged with **-ć-* or **-s-* already in Proto-Indo-Iranian. Perhaps this represents an enlargement with the substratum suffix **-sa-* (cf. LUBOTSKY 2001:304), that is: **xrausa-* ‘call, cry’ vs. **xrauća-* ‘caller, crier’.¹²

¹¹ It has generally been believed that the *-š-* in this Persian form comes from OIr. **-sġ-*, and it is often believed that OIr. **-sġ-* yields Pers. *-š-* (see for example HENNING 1933:207, who also cites previous literature). I take the opportunity here to demonstrate this is a misconception. All inherited examples show that OIr. **#sġ-* > MP, Pers. *#siy-* (as in *siyāh* ‘black’), and OIr. **-sġ-* > MP, Cl. Pers. *-h-* (plus palatalization of a preceding **a > i*, and regularly **-sġa-* > MP *-hī-*): OIr. **masġa* ‘bigger’ > MP *mahy-*, Cl. Pers. *mih* (< **mihī*), OIr. *kasġa* ‘smaller’ > Cl. Pers. *kih*, Proto-Persian **māsġaka-*, ‘fish’ > MP *māhīg*, Cl. Pers. *māhī*. Furthermore, this so-called sound law is not mentioned in SALEMANN (1895–1901). Now the three examples generally adduced to “prove” that OIr. **-sġ-* changed to Pers. *-š-* are NP *āš* ‘soup, nourishment’ < OIr. **āsġa-*, *kašaf* ‘turtle’ < OIr. **kasġapa-* and the verb MP *handēšīdan*, NP. *andēšīdan* ‘to think’. The verbal root *ās-* ‘to eat’ does not exist in Iranian (see fn. 17). The word *kašaf* ‘turtle’ can evidently not be inherited, since word-final *-p* does not become *-f* in MP or NP, but *-b*, so this obvious loanword cannot be used to prove any sound law (**kihab* would be the expected NP reflex). As to the *-š-* in MP *handēšīdan*, NP *andēšīdan*, I believe it is analogical of the subjunctive, imperative and injunctive moods, OIr. **daiš-* (cf. OAv. *dāiš*, etc.). This is made even more probable, in my opinion, in view of the semantics of this verb, which originally meant ‘to show, to demonstrate’, and with **ham-* ‘to be shown, to be told’ (cf. CHEUNG 2007:51–52). This also probably happened in Wakhi *diš-:dišt* ‘to know, understand, to be able’ (I thank Prof. KÜMMEL for pointing out the Wakhi parallel to me). HENNING (1933:207) proposes that this change OIr. **-sġ-* > Pers. *-š-* is due to a third, unknown dialect that would be one of the sources of Persian vocabulary. Nevertheless, given the scarcity of examples, most of which are clear borrowings, as well as the number of counter-examples, it seems safer to simply renounce this sound law and consider all examples of Pers. *š* against OIr. **sġ* either as borrowings or having a different origin than commonly assumed.

¹² Nevertheless, a borrowing cannot be excluded as the source of the NP forms, or, perhaps this form goes back to PIIr. **krauć(V)-* > PIr. **xrauxšt-* > Proto-Persian **xrōšt*. However, the absence of a final *-t* in this form makes this hypothesis weaker.

I would like to mention two other related forms, both found in Early and Classical New Persian: *xurōč* ‘rooster, cock’ and *xurōj* ‘(loud) call’ (STEINGASS 1892:457). If inherited, *xurōč* could perhaps go back to OIr. **xrauščV-*. It is also quite possible that these words were borrowed from another Iranian language, but I cannot tell which one. Their internal derivation seems very difficult.

The alternation between PIr. **xrauča-* (yielding most of the words cited above) and PIr. **xruća-* (yielding PG **hros*, Sistāni *xross*) ‘loud animal’ seems even more problematic to explain from within Iranian. If we take these forms, or the root in general, as being of substratal origin or a borrowing,¹³ perhaps this could – to some extent – explain the irregularity of some of the forms. In this case, perhaps we could explain Sanskrit *krauñca-*, m. ‘crane’ (also *kruñc-*, with the same meaning) similarly: also built on a PIr. **kru-* ~ **krau-* base with the addition of a “substratum” suffix *-ñc-* (which has never been satisfactorily explained otherwise, see HAMMER 2019:401–402 with references to the literature). This “suffix” has not been recognized as such, but the sequence *-ñc-* is found in a number of words without an Indo-European etymology in Indo-Iranian languages. Some examples are discussed in Section 9.4., to which we can also add the Sanskrit verbs *kruñc-* and *kuñc-* ‘to bend’, without a convincing Indo-European etymology, and Sanskrit *prakuñca-* m. ‘ein bestimmtes Hohlmaß’, also without etymology (EWAia³ 340–341). To conclude, the forms discussed, although seemingly regular, or apparently presenting few problems, are more than problematic from an Indo-Iranian, Indo-European point of view. Moreover, considering that they all derive from one or multiple borrowings, probably or possibly made at the Indo-Iranian level, or independently by Proto-Iranian and Proto-Indo-Aryan, would, to some extent, make the entire question easier to consider. Last but not least, there are no Indo-European cognates to the Indo-Iranian forms cited in this Section.

6. *kalpak* Kermān Gav. ‘مارمولک’, ‘small lizard’
(GHOLAMI & FARAHMAND 2016:105, 157),
Kermān Persian *kalpak* ‘id.’

6.1. Cognates

This word has many cognates, cf. VOSKANIAN & BOYAJIAN-SURENIANTS (2007:123), which I cite here exhaustively (see also 6.4.1.):

¹³ Cf. Proto-Mon-Khmer **kraw* ~ **krāw* ‘to call out, announce’ (SHORTO 2006:474). Of course all these roots can be onomatopoeic in origin, but that does not preclude them from having researchable etymologies.

“S[outh]B[a]škardi [...], N[orth]B[a]škardi [...] *kalpak*, Mīnābi *kalpak* < OIr. **karpaka*-; cf. YAv. *kahrpuna*- “Name eines daēvischen Tiers” [(BARTHOLOMAE 1904:455)], Khwar[ezmian] *krbwn* [karbun] “lizard”, and MPers. *klpwk* [karpōk] [...] denoting an *Ahrimanic* animal, which is derived by Nyberg from OIr. **karpawaka*-. The same OIr. base can be probably traced also in Lār. *kalpōk* “big lizard”, and Banāfi and Rīči *kalpūk* “lizard”. [...]”.

Bandari *kalpak* ‘lizard’ can also be added to this list.¹⁴ As one can see, the lexeme is widespread, ranging from the South and the West of Iran as far as historical Khwarezm, and of course encompassing Fārs and the region of Kermān.

6.2. Derivation of the forms

The Gavruni word *kalpak* is probably a borrowing from another Iranian dialect since *ṛrp* is expected in Gavruni, which developed the phoneme /l/ only secondarily (see Section 6.3. for the inherited Gavruni form). It is possible that MP <klpwk> also transcribed [kalbo:k] or [kalbu:k] and that it is a borrowing, either from a dialect, or from a nearby Iranian language.

From the forms presented above, three different proto-forms can be reconstructed for this lexeme: **karpuna*-, **karpauaka*- and **karpaka*-. In fact, the latter two reconstructions are *Transponaten*: they represent the suffixes -ōk and -ak on a Middle Iranian base **karp*- (or rather, here, **kalp*-) ‘lizard’ (pace NYBERG 1974:112), **kalp-ōk* designating big lizards and **kalp-ak* designating smaller ones, reflecting the diminutive function of *-ak. It is also possible that **kalpak* is older, and reflects an -a-ka- derivation on a thematic form of the root **kalp*-.

The Banāfi and Rīči form *kalpūk* is probably an older formation: it is a -ka- (*ka*-) suffixation on a base **karpū*-, and the Av. and Khwar. forms could go back to a -na- suffixation on the same base. It is also possible, as proposed by DE VAAN (2000:284; 2003:582), that **karpuna*- represents an *-una- derivation on a stem **karp*-. One could perhaps also compare either *-u-na- (or *-una-) with the ending -yan- ~ -un- found in Skt. *átharvan*- and Av. *āθrauuān*- / *aθaurun*- ‘priest’, also of likely substratal origin (LUBOTSKY 2001:310).

¹⁴ I thank Aida ALAVI (Bordeaux University) for giving me this form.

6.3. Inherited Gavruni form

Kerm. Gav. *kerpu* ‘سوسمار’ (SFB 122)¹⁵ < Pre-Proto-Gavruni **karpū-ka-* seems to be an inherited Gavruni formation of the same root. It is also possible that it was borrowed from another language at a very early stage.

6.4. Discussion of other forms

6.4.1. Forms suffixed with *-āsa-*

A number of derivatives of **karp(a)-* enlarged with the suffix *-āsa-* are found in Iranian languages. I will enumerate some of them. STEINGASS (1892:1021) provides *kirbāsa* ‘a lizard’,¹⁶ *karbāyis* کربایس and *karbāyiš* کربایش ‘a venomous kind of lizard’, *karbas* کربس ‘a large venomous lizard’, *karbasū* کرباسو ‘a poisonous lizard’, and the same variants are found with the phoneme /ʃ/: *karbaš* کربش ‘a large and poisonous lizard’, *karbašū* کربشو ‘id.’, *karbaša* کربشه ‘a lizard’. There is also *karba* کربه ‘a large kind of venomous lizard’. The same variants are also found with /rp/: *karpāsū* کرباسو, *karpāsa* کرباسه, *karpāšū* کرباشو, *karpāša* کرباشه, all designating ‘a sort of poisonous lizard’, and with /lp/: *kalbāsū* کلباسو ‘a large lizard’ (STEINGASS 1892:1042), Tajiki *kalpēsa* ‘lizard’, possibly of Sogdian origin (NOVÁK 2013:95), and, finally, Sogdian <krps’k> *karpasē* (see GHARIB 2004:193, who transcribes it wrongly as *karpāsē*). Given all the variation in vowel length, type of suffix, consonants, etc., one could postulate that either these forms were borrowed, or they underwent taboo deformations. DOERFER (1975:303) proposes both: he suggests that all the forms are from Turkish *kālpāzā* “und ähnlich.” “Im Tü. gibt es sehr viele Formen (wahrscheinlich Tabubildungen)[...]”. Since the *kālpāzā* form (or any similar form meaning ‘lizard’) is not attested in any ancient Turkic language, I do not see why this should be a borrowing from Turkish and not from Persian into Turkish. Also, Turk. *kālpāzā* would probably be rendered in Cl. Persian as **kilpizi* or **kalpezi/a*, but an Iranian form of this word with a /z/ has yet to be found.

Forms like *karbāyis* and *karbāyiš* کربایش which show *s/š* variation in a completely different part of the word than *karbaša* etc. could reflect a borrowing from a language where *s/š* are not phonemically differentiated. As no such

¹⁵ *sūsmār* سوسمار although etymologically meaning ‘worm-snake’ here probably designates bigger reptiles, such as varans and crocodiles. It could nevertheless also designate geckos and other types of lizards.

¹⁶ This word and many others have other meanings as well (usually due to homonymy), but I only cite the relevant meaning here.

donor language is known for Persian, it seems likely that we have here to resort to an explanation with taboo deformation which, given that all these lizards are venomous, is understandable. There is also باشو *bāšū* ‘a lizard’ (STEINGASS 1892:147) which is probably extracted from **karbāšū*.

The word چلیپاسه *čalpāsa* ‘a small venomous lizard’ (STEINGASS 1892:398) might reflect a synchronic variation **čarp-* ~ *karp-*, cf. Section 6.4.2. This is maybe linked to the *kar-* ~ *čar-* variation seen in, for example, MP *kargās*, NP *kargas* ‘vulture’ : Sogdian *čarkas*, ‘id.’, Oss. *cærgæs* ‘eagle’ (CHEUNG 2002:176), Sogd. *čarxušt* ‘wine-press’ : Pahl. *karxōš* ‘id.’ (but NP *čarxušt* ← Sogd., cf. HENNING 1939:96–97). The word *čalpāsa* is in any case a borrowing from another Iranian language into Persian, as are all of the Persian words cited above, since **karpāsa* should have yielded **karbāh* in NP, possibly found in MP <klb’h> (NYBERG 1974:112). Nevertheless, Sogdian *karpasē* ‘lizard’ does not have a variant with *čarp°* for this word, so Sogdian cannot be the source of NP *čalpāsa*, or, perhaps, the Sogdian form starting in **čarp°* (cf. Sarikoli *čarbost* ‘lizard’) was itself replaced by a loanword.

6.4.2. Discussion of the Ilr. suffix **-āčá-*

DE VAAN (2000) discusses the Indo-Iranian suffix **-āčá-* in the most comprehensive recent work on this suffix. His conclusion is that “[t]he Ilr. suffix **-āčá-* came to be used in Iranian to denote a larger variant of the animal to which the basic noun refers.” His examples are: “**karka-* ‘hen’ → **karkāčá-* ‘eagle’, **karpa-* ‘frog?’ → **karpāčá-* ‘lizard’ and Khot. *mura-* ‘bird’ → *murāsa* ‘peacock’ [...]” (p. 285) and “**(H)rupi-* ‘marten’ → **raupi-* ‘fox’ → **raupāčá-* ‘fox’/‘jackal’.” (p. 290). DE VAAN further claims that **-āčá-* was a foreign suffix, borrowed by the Indo-Iranians from a substratum language (p. 290–291). He rightly rejects the old etymology of **karkāčá-* as ‘hen-eater’ (p. 285–286).¹⁷ DE VAAN further notes (p. 279–80) that there are variants in some Iranian languages that seem to go back to **raupāčá-* with a short *a*,¹⁸ and

¹⁷ According to MAYRHOFER, the alleged Indo-Iranian root *aś-* ‘essen’ has only two known reflexes in Iranian: Av. *āsītō*, which cannot be related to it (see DE VAAN 2000:292⁹), and Av. *kahrkāsa-* “Geier < *Hühnerverschlinger” (EWAia¹ 136). As pointed out by DE VAAN (2000:285) vultures are not typical ‘hen-devourers’: as is known, they eat most kinds of carrion, but only some species attack (in some circumstances) live animals. In fact some American species can kill and eat hens, but even so, it only happens rarely. More importantly, vultures are seen as good animals in Zoroastrianism, and roosters are sacred: good, pure animals do not eat sacred creatures, and naming animals in an auspicious way is an important matter, cf. the controversy on the name of the animal-demon mentioned in Videvdad XIII:6.

¹⁸ To the forms cited (Sogdian, Khwarezmian, Munji, Ormuri, Šuyni and Yazgulami), one can add Gavruni *ruvas* (SFB 91) (but HOUTUM-SCHINDLER 1882:63 Kerm. Gav. “Fuchs *rūwās*” a

that variants of **karkā́ca-* and **karpā́ca-* “display a similar shortening”. Instead of a secondary shortening (which would have no explanation, as admitted by DE VAAN 2000:280), I believe the variation already existed in the oldest stages of (Proto-)Indo-Iranian. This could be an argument in favour of a foreign origin of this suffix, as suggested by DE VAAN (2000:290–291). The suffix is thus, in fact, **-ā́ca-*. A similar alternation is seen in PIr. or OIr. **kunčī́ta-* ~ **kunčī́ta-* ‘sesame’, also of foreign origin (see p. 54). An internal explanation of this alternation would be that, when added to a thematic base, the suffix **-ā́ca-* was lengthened due to a *sandhi*-effect, so, virtually: **karpa-* + *-ā́ca-* > **karpā́ca-*, and when added to the athematic base, or to a form ending in another vowel, **-ā́ca-* would be reflected as short: **karp-* or **karpu-* + **-ā́ca-* > **karpā́ca-*. Similarly, **raupi-* + **-ā́ca-* > **raupā́ca-* while **raupā* + **-ā́ca-* > **raupā́ca-*. In the case where this is correct, one could also resort to the same reasoning to explain the **-ita-* ~ **-ī́ta-* alternation: PIr. **kunč-* + *-ita-* > **kunčī́ta-* ~ **kunčī́-* + *-ita-* > **kunčī́ta-*. Another phonetic peculiarity relative to the **-ā́ca-* suffix is the existence, in many Iranian languages, of *ča-* variants in the root of both **karkā́ca-* and **karpā́ca-* (with unclear distribution), leading us to reconstruct **karkā́ca-* ~ **čarkā́ca-* and **karpā́ca-* ~ **čarpā́ca-*. An important point is that, as far as I know, there are no *ča-* variants of the unsuffixed base nouns **karka-* and **karpa-*.¹⁹ Although this too might point towards a foreign origin of the suffix, I am not quite sure how to explain it. If we posit that *-ā́ca-* provoked palatalization of the initial **k-*, as seems possible, yielding **čarpā́ca-*, **čarkā́ca-*, then forms such as Av. *kahrkā́sa-* were back-formed from the bare noun **karka-* (**karka-*: **čarkā́ca-* > **karka-*: **karkā́ca-*). On mere semantic grounds, I do not believe that the **-ā́ca-* suffix “denote[d] a larger variant of the animal to which the basic noun refer[ed].” (DE VAAN 2000:286). Whether PIr. **karkā́ca-* meant ‘vulture’ or, as is more likely, ‘eagle’, we would still have to justify why an eagle or a vulture would be called ‘a big hen’ and not, for instance, ‘a big bird’. Hens, eagles and vultures are very dissimilar in appearance – and their size is far from being their main difference. More problematic even is the claim that lizards are larger frogs. The meaning of ‘frog’, found in some languages for the word **karpa-* and its derivatives is by

form which, if extant, could always have been produced under the influence of Cont. Pers. *rubāh*).

¹⁹ Except in Pashto *čərg* ‘cock’, *čərga* f. ‘hen’, where it is a regular reflex of **kṛkṛja-* (MORGENSTIERNE 1927:19), and Ormuri *čirgōtai* m., *čirgōtiē* f. ‘chicken’, diminutives formed on Pashto loanwords (MORGENSTIERNE 1973:125).

all accounts secondary.²⁰ If anything, the existence of the meaning ‘frog’ for this word in some languages could imply that the initial, original meaning of **karpa-* was different from both ‘frog’ and ‘lizard’ (see Section 6.5.).²¹ Furthermore, the reflexes of **karpǎća-* cited by DE VAAN (2000:284) have the meaning of ‘frog’ which would imply that there was a development of meaning from *‘frog’ to *‘lizard’ and then to ‘frog’ again. This does not seem convincing to me. I also do not think that a peacock is simply a ‘larger bird’, and that a fox is a larger kind of fox as would be the natural conclusion to DE VAAN’s reasoning. In fact, I believe that this suffix indicates dangerousness, not only to humans, but also to flocks, herds and property.²² For instance, while the forms cited in Sections 6 and 6.1 designate smaller and bigger lizards, the forms cited in the Section 6.2.1, containing the suffix **-ǎća-* mostly designate venomous lizards. Peacocks are dangerous birds, they attack property, grounds, smaller animals and human children.²³ The **karkǎća-* would thus be a dangerous bird, which seems to confirm DE VAAN’s intuition that the original meaning was ‘eagle’ (2000:286). In any case it should be a preying bird (not necessarily large). The meaning of preying bird evolved to designate vultures in regions where those are more common, and are also, in multiple ways, preying birds. Finally, if **raupi-* designated martens, who are omnivorous but mostly harmless to humans and herds, then **raupǎća-* would have meant *‘dangerous type of marten’ > ‘fox’. The meaning ‘fox’ attributed to **raupi-* could have been, in a way, “back-formed” from **raupǎća-* ‘fox’. I believe such a morpheme implying the dangerousness of an animal must initially derive from a nominal composition process; only later did this morpheme become grammaticalized as a suffix. The original meaning of this morpheme could be something like ‘venom’, ‘snake’, ‘killer’, designate a mythical monster like ‘dragon’, or simply mean ‘monster’. This process is also found in living languages, for example in

²⁰ DE VAAN (2000:291⁷) suggests that, if the meaning ‘frog’ were primary, the word could derive from “the Ilr. root **krap-* ‘to lament, to wail’ [...]; the motive for the designation would have been the sound which frogs produce.” In fact, I do not understand the metathesis proposed here, nor do I understand the nominal derivation, as **krapa-/*karpa-* should then mean ‘lamenting’, not ‘the one who laments’. In any case, I believe ‘frog’ cannot be the primary meaning, cf. Section 5.4.

²¹ Martin SCHWARTZ suggested to me that this word could have initially designated a type of gecko, the *Teratoscincus*, common in Central Asia and Iran, and which bear resemblance to both lizards and frogs.

²² Note that this suffix is not added to the names of already dangerous animals such as wolves, panthers or lions.

²³ Cf. for example this article from the Houston press:

<https://www.houstonpress.com/news/peacock-attacks-actually-do-exist-6726626>
(14 February 2014).

English: *dragon-* in *dragonfly* so called because of its shape and possibly because it can bite and hurt human flesh (the dragonfly is also called the *devil's darning needle*), *go* 'cow' is used in compounds in Gavruni to designate bigger wild animals: *go-gurāz* 'wild boar' (lit. 'cow-boar'), *go-miš* 'buffalo' (lit. 'cow-sheep') (HOUTUM-SCHINDLER 1882:63). In another way, *gok* 'frog' is used to determine *som* 'the digger' in *gok som* 'mole cricket' (see p. 31). It does not seem absurd that a compound with a word like 'snake' or 'dragon' would give the sense of 'dangerous' to an animal name. This second member of the compound would later be grammaticalized as a suffix. Naturally, this explanation remains speculative.

Lastly, if, as proposed by DE VAAN (2000:290), the suffix **-ā́ca-* has a variant **-ā́j^há-* (thus probably **-ā́j^há-*) as found in **uarā́j^há-* 'wild boar' (a substratum word, cf. LUBOTSKY 2001:303), then **uarā́j^há-* could mean 'dangerous boar' < **uara-* 'boar, pig' (of which we have no attestation). This is also very speculative, as are the identity of both suffixes, and the analysis of **uarā́j^há-* as **uar-ā́j^há-* (< **uara-aj^há-*?).

6.5. Etymology of **karp-* ~ **kalpu-*

The meaning of **karp-* and its variants **kalpu-*, **karpu-*, **karp-*, **kalp-* is not exactly 'lizard' but, since it means either 'frog', 'lizard' or 'demonic creature' in the Old, Middle and New Iranian languages, it seems to me that it rather designates a despicable creature of the type of lizards and frogs. It is the prototype of the Zoroastrian *xrafstra-*, a creature of the Demonic Spirit, that hinders the furthering of the good creation.

There are two possible etymologies: either **karpu-* etc. originally corresponds to a substratum root, borrowed by Proto-Iranian (or by Proto-Indo-Iranian but subsequently lost in the Indian branch), or it is inherited from Proto-Indo-European. In the first case, it seems tempting to posit a proto-form **kalpu-* (with a secondary variant **kalpa-*), and to explain the forms with **/rp/* (e.g. Avestan, Khwarezmian, Sarikoli, Gavruni, some Persian forms) as being adaptations of **kalpu-* in those languages which, at some point, did not have the phoneme */l/*. In the second case, the inherited sequence **-rp-* could have been dissimilated to **-lp-*, already in the Old Iranian stage. Both hypotheses suppose the existence of */lp/* ~ */rp/* variants in this series of words at the Old Iranian stage, perhaps even in Proto-Iranian.

It is possible that **kalpu-* is a cognate of Av. *xrafstra-* 'obnoxious creature', which has no good Indo-European etymology. BAILEY (1970:27) proposes an analysis of *xrafstra* as *xraf-* "with quadruple suffix *-s-t-r-a-* (or as a verbal

xrafs- with IE *-s-* enlargement [...]).” From a morphological point of view, neither proposals are sensible. BAILEY’s attempts to connect *xrafstra-* to PIE **(s)kerp-* and the Indo-European names of the scorpion (1970:27–28) are not methodologically sound. Nonetheless, his proposal (1970:28) to connect *xrafstra-* to **karpu-*, *kalpak*, etc. seems convincing, on the semantic basis that both designate a ‘disgusting, repugnant animal’. An Indo-European etymology should nevertheless be abandoned for these words, which, in my opinion, derive from a substratum root **karp-* ~ **krap-* or **kalp-* ~ **klap-* ‘disgusting, repugnant animal’.²⁴

A number of wild animal names of substratal origin have been noted by LUBOTSKY (2001:307). To further corroborate the substratal hypothesis concerning the origin of these words, as Professor Lubotsky suggested to me, the *-stra-* in *xrafstra-* looks similar to the **-stra-* cluster in **Huštra-* ‘camel’, also a substratum word (LUBOTSKY 2001:307). If it is the same suffix, there is one chronological problem: why did **Huštra-* undergo the ruki change of **us > uš* while **krapstra-* did not undergo the Avestan sound law **ps > fš* (HOFFMANN & FORSSMAN 2004:105)? To this question, all possible answers I can think of are speculative: perhaps *xrafstra-* was borrowed into Avestan after the sound law ceased to be operative? Perhaps the proto-form of this word was **krafstra-*, and **-fs-* (only found in loanwords) did not become *fš*? Or perhaps, more straightforwardly, did **-fštr-* dissimilate to *-fstr-* (**-fšt-* dissimilates to *-fš-*, cf. examples in HOFFMANN & FORSSMAN 2004:105). Another possibility, mentioned to me by Dr. Michaël PEYROT (p.c.) is that *xrafstra-* goes back to **xrafttra-* (which would regularly become *xrafstra-*). This is possible, but implies that the suffix here is *-tra-*, not **-stra-*. In that case, **Huštra-* should probably be analyzed as **Huš-* + *-tra-*.

²⁴ HENNING (1947:41¹) proposes that Manichean Middle Persian ⟨prystr⟩ /frestar/ ‘noxious creature, reptile’ is “a corruption of Avestan *xrafstra-*”. GERSHEVITCH (1954:246) suggests the opposite: **frafstra-* “may still happen to have been the original form from which Av. *xrafstra-* was dissimilated.” He further suggests to derive it from **fra-pt-tra-*, *pat-* describing the movements of *daēvic* beings in multiple texts. This etymology is generally accepted by current scholarship, but it faces several problems, besides being *ad hoc*. First, the vowel noted by ⟨y⟩ in the Manichean Middle Persian script might render *ī*, *ē* or *ĩ*, *ẽ* (as a *mater lectionis*). This is not an expected reflex of **fra-*, cp. Old Persian *fravarti-* > Middle Persian *frawahr*, OP *framānā-* ‘command’ > MP *framān* ‘id.’. Second, the suffix *-tra-* is completely obscure here: since the word needs to be inherited, in GERSHEVITCH’s scenario, it has to be the instrumental suffix *-tra-*, which does not function semantically here. Third, the root **pat-* has the meanings ‘to fly, to jump’ or ‘to fall, to stumble’ (hence the *daēvic* use of this verb), while *xrafstra-* seem to originally be creeping creatures, or earthly creatures.

6.6. Etymology of Av. *karapan* ‘evil priest’

Finally, I will take advantage of the present etymological discussion to add another hypothesis to the lot: the (Old and Young) Avestan word *karapan-* ‘evil priest, person who neither listens to nor sees the True Teachings’, nom. sg. *karapā*, nom. pl. *karapanō*, gen. pl. *karafnqm*, does not have a widely accepted etymology. BARTHOLOMAE (1904:455) proposed to link it to Ved. *kālpa-* ‘rite’. However, this root does not exist in Iranian. ABAEV (1956:53) proposed to connect this word to Ossetic *kæraef*, *xælaef* ‘алчный (= greedy)’, Hittite *karap-* ‘пожирать (= to devour)’, and *xrafstra-* (< **k(a)rap-s + -tra*). This does not hold semantically: *karapan* are not devourers, nor are they particularly greedy, in the sense of desiring material goods. The etymology of *xrafstra-* from **k(a)rap-s + -tra*, if inherited, does not work (cf. Section 6.5). Various other proposals have been put forward (for those published up to 2010, see MALANDRA 2010). Much more convincingly (both formally and semantically), SCHWARTZ (2013:64) proposed to link this word to an Indo-Iranian stem **kṛpá-* ‘make a plaintive sound, implore’, reflected by Khwarezmian *krb-* (/kirba/) ‘to moan, to mumble’. He supports the connection of this root to the notion of ‘priest’ with a series of Vedic parallels.

This word has no known cognate in any other Iranian language, nor in Indic, and seemingly no longer represented a reality by the time of Young Avestan, cf. MALANDRA 2010: “[r]estricted as they are to only formulaic status, the *karpans* had become only a distant memory by the time they are mentioned in the *Yašts*.” It seems more likely that the word *karapan*, although related to SCHWARTZ’s reconstructed root, is an innovation of Avestan, rather than having been lost in every other single Iranian branch and language.

There are two major problems concerning the word *karapan-*, one concerning the suffix, and the other concerning the root.

6.6.1. Av. *karapan-*

As DE VAAN (2003:130) writes: “For a few stems, it is unclear whether we must reconstruct *-ān-* or *-an-*: [t]he etymology of OAv. *kar[a]pan-* ‘karapan, désignation d’adversaires religieux’ is uncertain, so that it is unclear whether the nom. pl. *karapanō* [...] has undergone shortening. If the word represents **kalpa-Hn-* ‘pertaining to arrangements’ (to Skt. *kālpate*), we would expect †*karapānō*, but a stem **kalp-an-* is also conceivable. [...]”²⁵ In fact, this poses

²⁵ Elsewhere in the same book (2003:581) DE VAAN seems convinced by this etymology: “it was connected with Skt. *kālpa* ‘ritus’ by BARTHOLOMAE (1904: 454–455), from which we can now

a problem: as DE VAAN (2003:468) states, the regular reflex of Plr. **-anV-* is OAv. *-ān-*. All other cases, he writes “will be due to restoration of *-an-* on the basis of the YAv. form.” He then cites a number of forms, among which *karapan-*. This is, in my opinion, quite difficult to believe: only the genitive plural, *karafnqm*, is attested in Young Avestan, and then, only in fixed formulas. Verily, the notion of *karapan-* is a Gathic one. Why would the scribes restore *-an-* in Old Avestan on the basis of a Young Avestan form that barely existed, for a concept that is virtually unarticulated in the entire extent of Young Avestan literature? BEEKES’s explanation (BEEKES 1988:61) that “[t]his word has *-an-*, not *-ān-* (< **-on-*) in the oblique cases, and these words had suffix accent”, implying a leveling from the oblique cases, is more fitting.

6.6.2. Av. *karapan*

The second problem linked to this word is the accent. “The noun *karapan-* is shown by the Gathic metre to count as disyllabic /karpan-/” (DE VAAN 2003: 581). As is known, if the word were stressed *kárpan*, it should be spelled **kahrpan* which is why BEEKES (1988:61) says “these words had suffix accent”.²⁶ DE VAAN (2003:582) has a slightly different solution: “[f]or some reason, original **karəpan-* was changed to *karapan-* in the history of our texts. If this has occurred before VOR [Voicing Opposition on **r*, my note], the noun *karapan-* would be irrelevant for the present discussion. If the change occurred after VOR, we would need to assume that the nom.sg. **kárpā(n)* adopted the accentuation or at least the voiced variant *kar-* of the oblique cases. [...]”. I believe that the sound law *árp > ahrp* cannot have been blocked by the secondary <a>, as DE VAAN seems to propose, since the second <a> in this word is purely graphic: the word is doubtlessly syllabified as /kar.pan/. If the word is to be reconstructed as **k(a)rp-h₁en-*, a suffix accent is unlikely. We thus have a word /karpan-/, which cannot have suffix accent, but also cannot have an old root accent. If we accept its derivation from an Indo-Iranian stem **k_ṛpa-*, as proposed by SCHWARTZ (2013:64), then we must explain its structure /karpan/ instead of **kərapan-* /kṛpan-/.

derive it satisfactorily by means of the individualizing suffix **-h₁n-* as **kalpa-Hn-* [...]”. It is of course difficult to avoid self-contradiction in such a huge volume.

²⁶ If not a typographical mistake he also includes *kahrpuna-*, which obviously cannot have suffix accent.

A convincing solution to this problem was suggested to me by Martin SCHWARTZ.²⁷ One has to suppose that the Indo-Iranian root **kǵp-* ‘make a plaintive sound, implore, to complain’ (cf. EWAia¹ 409),²⁸ (1) was thematicized as **kǵpá-* (adjective) ‘imploring, complaining’, then (2) was substantivized with the *Catō*-suffix **-n-* (noun) **kǵpán-* ‘implorer, mumblor’, yielding **kǵpán-* in Proto-Avestan (like **marta-* ‘mortal’ → *martān-* written as <marətān> ‘mortal man’) which was then influenced by the root **karp-* ‘disgusting, repugnant animal’. This later step changed **kǵpán-* into **karpán*,²⁹ which would naturally yield **<karapān>* and would therefore explain why we have *-an-* and not *-ān-* (cf. Section 6.6.1.). The observed shortening is difficult to explain, but one could perhaps resort to an analogy with *kauui-* ‘(evil) ruler’, a word constantly co-occurring with *karapan-* throughout the Avesta (as suggested by Prof. WEISS, p.c.). I believe that the oblique stem of *kauui*, *kauuai-*, is originally short (cp. Lydian *kawes* ‘priest’ < PIE **kouh₁-é₂-*), but in the case that it is originally long (as proposed by Prof. SCHWARTZ, p.c.) the analogy would function better. The second graphic <a> in <karapan> was added to indicate that the word should not be pronounced /kahrpan/, exactly like Av. <marakaēcā> /markájt̪ša/ (Yasna 31:18) vs. <mahrka> /márka/ ‘death’, cf. HOFFMANN & FORSSMAN 2004: 117f.³⁰ This title would thus be explicitly insulting to this category of priests, which the Prophet Zarathushtra despised. In this way, the fact that *karapan-* is a purely Avestan word, absent from the rest of Indo-Iranian, is also accounted for.

**7. *karatin* Yazd Gav., *keratin* Kermān Gav. ‘spider’ (various sources),
kartonak Tehran Gavruni (SFB), cf. Meymei *karatena* ‘id.’**

7.1. Etymology

This word, already cited in IVANOW (1939:125), is still in use. According to IVANOW (id.) “[t]he meaning of *kara* is not clear, but *tin* is the Pres. stem from *tinūdwun*, to weave”. IVANOW’s intuition seems to be confirmed by the structure

²⁷ I wholeheartedly thank Prof. Martin SCHWARTZ, since this entire discussion stems from an email exchange with him. All the conclusions I present in this section are also his, but all remaining mistakes are mine alone. He also kindly asked Prof. Michael WEISS, whom I equally thank, for details concerning the formation of *karapan*.

²⁸ This root itself could be of foreign origin in Indo-Iranian, as it lacks any Indo-European cognates (EWAia¹ 409). Prof. Martin SCHWARTZ (p.c.) suggests it is of onomatopoeic origin.

²⁹ In an email dated to the 16th of November 2019 Prof. Martin SCHWARTZ writes: “I am now inclined to believe that the original noun had a syllabic *r* as do the Vedic comparands (cf. myself 2013 and Jamison 2012), and that the vocalism changed within Iranian. [...] I can think of two motivations: the binomial coupling with *kauai-* and an intentional pejorizing deformation [...]”.

³⁰ I thank Prof. Martin SCHWARTZ for pointing out this example to me.

of the word: an inherited (mono-morphemic) Proto-Gavruni **karatin-* (= *karátin* or *karatín*) should yield Yazd Gav. †*kratin*.³¹ If the first element is analyzed as a separate morpheme, the structure of the word is no longer problematic. In fact *kara* derives from *kara* ‘spider web’, otherwise unattested in Gavruni, cf. Cl. Pers. *kara* ‘spider web’, Yazdi Persian *kare* ‘id.’, sometimes compared with Yazdi Persian *kār bāftan* ‘to weave a tissue’, Cl. Pers. *kārtanak*, *kārbāfū* ‘spider’. EILERS, cited by CHEUNG (2007:244–245), suggests, on the basis of Pers. *kārtan* ‘spider’ a possible connection to Sanskrit *kart-* ‘to spin’ and the PIIr. root **kart-* ‘to twist, turn’. On the basis of Gav. *karatin*, Meymei *karatena* ‘spider’ and Qohrudi *kāre* ‘spider web’ (for which see CHEUNG 2007:245), this etymology should be completely abandoned, as sensed by CHEUNG (id.).

The words *kār*, *kāra* all relating to weaving, tissues and spiders have not yet received an etymology. My proposal is that they derive from *kār* ‘work’, implying that weaving is the work *par excellence*. This finds an exact parallel in French, where *ouvrage* ‘work’, was quite often used in the sense of ‘weaving work’ in the past, for example in the famous verses by BOILEAU DESPRÉAUX ([1674] 1825:18): “Hâtez-vous lentement et sans perdre courage,/Vingt fois sur le métier remettez votre ouvrage” that is “Hurry slowly and without losing heart/Twenty times again put your weaving back on the loom” (although he is metaphorically referring to the work of writing). While the word *ouvrage* very often meant ‘a weaving’ among the people, in more intellectual circles it meant ‘book’, as in the work *par excellence*, and it is still one of its more common meanings in French. Dutch offers another attractive parallel: *werkje* literally ‘little work’ designates either patrons or motifs of weaving, or ‘needle-work’.³²

In this sense, Pers. *kār bāftan* is an exact etymological parallel to French *tisser l’ouvrage*. Qohrudi *kāre* ‘spider web’ < **kārak* ‘small work’ seems to be parallel to Dutch *werkje*, with a semantic shift **‘needle-work’* > ‘web’. As to *kara*, *kare* ‘web’, they remain without an etymology for the moment. I believe they might go back to *kar-* ‘to do, make, build, shape’, cf. Old Persian *patikara-* ‘picture, (sculptured) likeness’ (KENT 1953:194–195) > MP *pahikar* ‘picture, image’ (CPD 63), NP *paykar* ‘statue’. Since in Old Persian *patikara-* the morpheme *kara-* “[has the] passive meaning ‘thing made’” (KENT 1953: 179), perhaps here too *kara* < OIr. **karaka-* had the original meaning of ‘small shaped thing, small thing made’ or, like in Dutch, ‘small thing done, small

³¹ The Kerm. form is ambiguous, and does not help with the reconstruction of the Proto-Gavruni form of this word.

³² I thank Prof. LUBOTSKY for informing me of the Dutch parallel.

work’ > ‘needle-work’ > ‘web’.³³ It is nevertheless also possible that this *kara-*, relatively restricted geographically, has a different etymology altogether.

7.2 Other forms

The Tehran Gav. word is an *-ak* derivative of a noun **karton* which I cannot explain, but whenever it differs from Yazd or Kermān Gavruni, Tehrān Gavruni often has somewhat aberrant forms, possibly taken from other languages, or modified by the speakers.

The Meymei dialect form *karatena*, an **-ak* derivative of **kara-ten-* (maybe **kara-tēn?*) must have the same etymology as the Gavruni forms.

7.3. Typological parallels: naming the ‘spider’

Calling the spider by its common action, spinning, is widespread among languages, see for instance Khotanese *vīśūna* ‘spider’, a derivative of “**uei-*” ‘to spin’ (DKS 388), Tocharian B *yape** < **ueb^h-o* ‘the spinner’ (ADAMS 2013: 520), Dutch *spin*, an ancient derivative of *spinnen* ‘to spin’ (EWN s.v. *spin*). The Arabic word *ṣankabūt* (with no established etymology) replaced the native Iranian words for ‘spider’ in many Iranian languages (including Persian, where it is much more common than *kārbāfu*), and its non-existence in Gavruni is, to some extent, surprising.

In Baxtiari, the spider is called *šaytoṽ* (ANONBY & ASADI 2014:196) from the originally Arabic word *šayṭān* ‘Devil’. This is particularly interesting in light of the fact that spiders are considered sacred, to some extent, in Islam. This way of naming spiders reminds us of, for example, the Purepecha word for spider, *xikwa-pu* ([ʃikwapu]), a derivative of the root *xikwa-* ‘referring to witchcraft’ (BELLAMY 2018:31, 255). Still in the American territory, it can be noted that in some languages of the North American plains, the word for spider derives from the name of a trickster character. For example in Arapaho *nih’óóθo* ‘Nihancan (a trickster character), white man, spider’, which is a cognate of the Cree hero name *Wīśahkēcāhk*, cf. GODDARD (1974:107–108).¹⁷ In a diametrically opposed way to the Baxtiari usage, the name *Hichaba Nihancan* (or *Hixcéébe Nih’óóθo*), literally meaning ‘(the) spider above’, designates God the Creator – the Heavenly Spider – in the Arapaho language.³⁴

³³ Perhaps in comparison with agricultural work, which was the “greater work”?

³⁴ Cf. the website <http://www.native-languages.org/arapaho-legends.htm>

8. *karpura* Gav. of Kermān ‘مریم نخودی’ = ‘(felty) germander’ (SFB), *kalpure* Persian of Kermān ‘id.’

8.1 Botanical identification of *karpura* and *maryam-e noxowdi*

The Kermāni word *karpura* is translated as مریم نخودی by Soruṣyān who gives the following definition: “گیاهی است کوهستانی و تلخ مزه با بوی تند که برای “اسهال و دل درد مفید است” (SFB 122): “*maryam-e noxowdi*. It is a mountain plant of bitter taste and spicy smell that is suited for (curing) diarrhea and stomach ache.” Since the dialectal (local) Persian word *kalpure* receives a similar description in DEHXODĀ’s dictionary³⁵ (except that, instead of mountain plant, it is called a desert plant), it is probably the same plant. They are also formal correspondents.

The compound مریم نخودی, *maryam-e noxowdi*, literally ‘pea-coloured (=beige) *Maryam*’³⁶ appears, according to some websites, to designate german- ders in general.³⁷ According to GOLFAKHRABADI & al. (2015), *maryam-e noxowdi-ye jangali* (‘sylvan *maryam-e noxodi*’) designates *Teucrium hyrcanicum*, in English ‘Iranian germander’. This seems to indicate that *maryam-e noxowdi* means ‘germander, *Teucrium*’. In another source, SCHLIMMER (1874:541), *maryam-e noxowdi* translates *Teucrium scordiodes* or “*germandrée aquatique*: Eng. *water-germander*.”³⁸ It is possible that *maryam-e noxowdi* designates different species of germander according to different people in different regions.

The plant *karpura*, *kalpure* cannot be ‘wall germander (*Teucrium chamae- drys*)’ (also found in Iran and in Persian called *xāmādaryūs* < Gr. χαμαίδρυς, cf. STEINGASS 1892:442 or *balūt el-arz* < Ar. بلوط الارض, cf. SCHLIMMER 1874:541), since wall germander does not grow in high places. *Teucrium hyrcanicum* should also be removed from the options, since it grows mostly in the Caspian region. I could not find any information on the smell of either plant.

Technically *karpura* could either designate felty germander, *Teucrium polium*, or *Teucrium marum*, cat-thyme. Both have a strong smell. *Teucrium polium* was traditionally used in food, at least in the Mediterranean region (but nothing indicates that *karpura* or *kalpure* are used as condiments). I believe that *kalpure* designates felty germander or *Teucrium polium*. Farsani, for instance, equates it with Baxtiari čez ‘pouliot de montagne, (N. Sci. *Teucrium polium*)’

³⁵ <https://www.vajehyab.com/dehkhoda/> کلپوره accessed on the 17th of July 2019.

³⁶ Note that SCHLIMMER (1874:514) transcribes it as «mèryèm nekhodi», that is /maryam naxodi/, which is an alternative pronunciation.

³⁷ According to various websites, including Wikipedia.

³⁸ Apparently called *suqūrdiyūn* in literary Persian (STEINGASS 1892:687).

(FARSANI 2011).³⁹ In KHADIGE & al. (2017), *Teucrium polium* is translated by *kalpure*. The traditional medical uses of *Teucrium polium* correspond more or less to those of *kalpure* and *karpura*.⁴⁰ It has also recently been discovered (KHADIGE & al. 2017) that *Teucrium polium* helps decrease the duration of the menstruation periods.

8.2. Etymology of *maryam-e noxowdi*

If I can propose an etymology for *maryam-e noxowdi*, I would like to suggest that it is related to *Teucrium marum* ‘cat-thyme’, called in Contemporary Persian *marv-e xošbu* according to SCHLIMMER (1874:541). According to DEHXODĀ, Cl. Pers. *marw* designates a type of *Lamiaceae*, a sort of basil, and it is equated to Ar. حبق الشيوخ the ‘pennyroyal of the sheikhs’ (as it translates) and Ar. ريحان الشيوخ, literally the ‘basil of the sheikhs’ which, according to LANE (1865:503), are (*Teucrium*) *marum*, cf. also Ar. المرؤ *al-marw* (id.). The Latin *marum* probably derives from Ancient Greek μάρον. The origin of one or the other form is obscure, but the Persian form *marw* was doubtlessly borrowed from Arabic. I believe that *maryam-e noxowdi* derives from a deformation of Lat. *marum* which was trivialized and made to sound more familiar to the Iranian ear and thus became *maryam*. It was then either generalized to mean ‘germander’ or specialized, designating other specific types of germander.

8.3. Etymology of *karpura* and *kalpure*

8.3.1. Immediate reconstruction

The word *karpura* needs to come from (Pre-)Proto-Gavruni **karpūrak*, that is, if it is inherited, Proto-Iranian **karpūraka-*. Since it appears that a number of local forms undergo *rp > lp*, and there is no clear example of *lp > rp* in Gavruni as far as I am aware, it is more than probable that the local Persian form *kalpure*, if it were inherited, also goes back to Proto-Iranian **karpūraka-*.

³⁹ This part of FARSANI’s PhD thesis has no page numeration, therefore I cannot cite the page number.

⁴⁰ The website *attarak* (<https://attarak.com/> گیاه-کلپوره accessed on the 12th of July 2019), for instance, lists a number of diseases boiled felty germander (probably in the form of herbal tea) is supposed to prevent or heal: constipation, cholera, colds, some fevers, liver dysfunctions, epilepsy, etc. It also prolongs life and helps a pregnant woman deliver painlessly.

8.3.2. In search of cognates

8.3.2.1. Sanskrit *karpūra*- ‘camphor’

Looking at similar words outside of Iranian, one finds the famous Sanskrit word *karpūra*-, masculine or neuter ‘camphor, camphor tree’ (EWAia³ 68) “wohl Fremdwort”. This word has a number of cognates which start in *kam*- ~ *ka*:- *camphor*, Dutch *kamfer*, Russian *ка́мфора*, Arabic and Persian *kāfūr*, Modern Greek *κάμφορα* ~ *καφουρά*.⁴¹

In Middle Iranian languages there is MP <k’pwl> [kāp/fūr] CPD 49, Sogdian <kp’wr>, <kpwr> [kapūr], GHARIB 2004:191, Khot. *kapūra*. The word *kāfūr* and related forms (also Hebrew *kopher*) either go back to Sanskrit (via Prakrit) or to Indonesian *kapur* (DONKIN 1999:81). If from Indonesian, it is due to “early Sassanian (226–641 C.E.) mercantile trade with South East Asia”. I personally favour this second hypothesis, since I do not see why the geminate (cf. Prakrit *kappūra*) should not have been preserved in any of those languages. The etymology of all of these forms is difficult, and has been much debated: ultimately, no consensus has been reached on the etymology of Sanskrit *karpūra*-, apart from that it is a loanword.

Two of the main types of camphor originate from South-East Asia (especially Sumatra, Borneo and Java), and the third from Eastern China, Taiwan and the south of Japan (DONKIN 1999:11, 50–68), before it was exported to the rest of the world. Although we find *karpura* ‘camphor’ in Old Javanese, it has been assumed that *karpūra*, derives from (Indonesian) *kapur* (DONKIN 1999:84).

GONDA (1932:23), explains that “[...] *kar*- occurs alongside *ka*- as an Austro-Asian prefix. This word also is thus Austro-Asiatic and not Aryan!”.⁴² GONDA’s proposal accounts for the variation between *kar*- and *ka*- forms, but not forms with *kam*-. This variation could in fact be explained, by comparison with the Proto-Mon-Khmer reconstructed form **knpur* ‘lime’ (SHORTO 2006:432), cf. Old Khmer *kapur* ~ *kampur* ‘chalk, lime’ (POU [1992] 2004:82, 85-6), Middle Mon *gapuiw*, Kuy *mphhò:r* (SHORTO 2006:432). In Old Khmer, there is also *karpura* ‘camphor’ (POU [1992] 2004:95), probably a borrowing from Sanskrit.⁴³

⁴¹ An account of words for camphor in a great number of languages is found in DONKIN (1999: 80–87).

⁴² “[...] *kar*- naast *ka*- als Austro-Aziatisch praefix voorkomt. Ook dit woord is dus Austrisch en niet Arisch!”

⁴³ According to SHORTO (2006:432), the Indonesian data suppose a borrowing from an Indonesian language into Proto-Mon-Khmer “with secondary infixation”, cf. Proto-Malayo-Polynesian **kapuR* ‘lime, calcium carbonate’ (BLUST & TRUSSELS, ACD), while there is a doublet **qapuR* in

The form with *kar-* could be explained by the infix *-rn-*, found in a number of Mon-Khmer languages (**karnpur* becoming **karpur* and **kampur* depending on the languages), cf. FERLUS (1977), but since only Sanskrit assuredly has an old form with *kar-* (all other forms with *kar-* being possible borrowings from Sanskrit), it is possible to simply explain the *-r-* here as a properly Indian phenomenon (cf. Section 8.3.2.3.).

Another argument in favour of a Mon-Khmer origin of this word is seen in the semantics. Prepared camphor is almost identical in appearance to lime and chalk, but lime and chalk are a much more common product, with a much more basic meaning in most languages. It seems to me that it is a more convincing scenario, to imagine a people who discover camphor and name it after lime, than them naming lime and chalk after a less common mercantile commodity. Gonda's proposal is very acceptable, and should be received, I believe, with the precision that this particular word comes from the Mon-Khmer branch of Austro-Asiatic. When exactly this borrowing occurred is unclear, and it could have happened at multiple times in multiple languages.

If the argumentation above is accepted, then Sanskrit *karpūra-* initially designates camphor as a product, and not the camphor tree or any plant. It is thus linked neither to PIr. **karpūraka-* 'germander' nor to Sanskrit *karpūraka-*, 'Curcuma Zerumbet', cf. below.

8.3.2.2. Sanskrit *karpūraka-* 'Curcuma Zerumbet'

Sanskrit *karpūraka-* m. 'Curcuma (or Zingiber) Zerumbet' (lex.), also called 'bitter ginger' is an aromatic plant, with fragrant leaves, used in food and for medicinal purposes. Its use is thus similar, at least in part, to that of the Gav. *karpura*, apart from the fact it grows in tropical Asia.

According to MAYRHOFER, Sanskrit *karpūraka-* is connected to *karcūra-* 'Gelbwurz' (EWAia³ 66), possibly having undergone contamination with *karpūra-* 'camphor', but since *karpūra-* is not a plant, the contamination hypothesis needs to be abandoned.⁴⁴

Perhaps one ought to reconstruct a PIr. **karpūraka-* 'medicinal plant with a strong smell'. It is unclear whether this word is formed by a double suffixation *-ūr-aka-*, whether it was morphologically **karpū-ra-ka-* or whether it was

Proto-Austronesian (id.). This naturally only partially concerns our data, since forms with *kam-* and probably Sanskrit *karpūra-* need to derive, in my opinion, from a Mon-Khmer language.

⁴⁴ Another argument against the identification of *karpūra-* with the camphor tree, or with any plant, is, from the *Hitopadeśa*, the name of the lake *karpūragaura*, which means 'camphor-like white'. The camphor tree is not white!

borrowed as such, with *-(ūra)ka-* being part of its root. There is perhaps yet another cognate in Gr. *κάρπιον* (lex.), cf. below.

8.3.2.3. Greek *κάρπιον* ‘*Pandanus odoratissimus*’

Found as a hapax in Ctesias, *κάρπιον* has been identified as the tree *Pandanus odoratissimus* (AMIGUES 2012:306), which, as its name indicates, smells quite strongly. As it is an Hellenized form of an Iranian original (AMIGUES id.), *κάρπιον* doubtlessly represents an Iranian **karpi*.⁴⁵ If its resemblance to **karpūra-* is not due to coincidence, then it is perhaps an *-i* derivative of a stem **karp-*. I do not understand further developments, but it seems that an Indo-Iranian, or separately Indian and Iranian root **karp-* with the meaning ‘odourful (medicinal) plant’ could be reconstructed. With this kept in mind, it is possible to explain Sanskrit *karpūra-* ‘camphor’ instead of expected **kampūra-* because of a contamination from the root **karp-*, and the word *karpūra-*, (that is, in the other direction than the one MAYRHOFER suggested), due to the odorous aspect of camphor. This, nevertheless, implies an antiquity to the root **karp-* or to the form *karpūra-* that is still impossible to ascertain.

9. *konjet*, *konjed* Yazd Gav., *kunjed* Šarpavoi (Boyce), *konji* Kermān Gav. ‘sesame’ (SFB), cf. Cl. Persian *kunjid* (Contemporary NP *konjed*) ‘id.’

9.1. Gavruni forms

The three different Gavruni forms imply different origins of the word in Gavruni dialects. One could reasonably argue that Šarp. *kunjed* was borrowed from Persian at a time when the historically short /u/ was still realized as [u] (as opposed to /o/ in Contemporary New Persian), and *konjed* was borrowed at a later period, while *konjet* is either the Gavruni inherited word, as can be deduced by the retention of the /t/, or a borrowing from Yazdi Persian.⁴⁶

The Kerm. final *-i* probably derives from a variant **/kundʒi:t/*, cf. Cl. Persian *kunjid*, a variant of *kunjid* ‘id.’, see Sections 6.4.2. and 9.4. Formally, it could also derive from **/kundʒi:k/*, but no evidence supports this reconstruction.

⁴⁵ TYCHSEN’s explanation (apud AMIGUES 2012:306) of *κάρπιον* by Pers. *kār-būy* lit. ‘make-smell’ is formally impossible: *kār* ‘*faciens*’ is a last-member of compounds, and in Ctesias’s times (Vth – IVth century BCE), ‘smell’ was certainly closer to PIr. **baūda-*, Av. *baōda-* than to contemporary Pers. *bu*, *buy*, or even Cl. Pers. *bōy*! Furthermore, how could *bōy* or *buy* be rendered by Greek *-πι-*?

⁴⁶ Although I have not documented this word in Yazdi Persian, it is very likely to exist under this form, since Yazdi Persian devoices word final *-d*.

9.2. Iranian cognates of Gav. *konjet*

Among the Iranian cognates of this word are Middle Persian <kwnc(y)t> ‘sesame’ (CPD 52),⁴⁷ Cl. Persian *kunjd* (Contemporary NP *konjed*) ‘id.’⁴⁸, Khotanese *kumjsata-* (DKS 61–2), Khwarezmian <kncd> ‘id.’, Baxtiari *konjet* ‘sesame seed’ (ANONBY & ASADI 2014:200). There is also Buddhist Sogdian <kwyšt’yc> ‘related to sesame or hemp’, a derivative of unattested *kuišt*^o, which GHARIB (2004:202) explains as coming from **kuncīt* > **kūinct* > **kūinšt* > *kuišt*, while BENVENISTE (1940:180) wonders whether it is “une mauvaise graphie pour **kwnšt* - ?”. There is also Pashto *kunjōla* ‘id.’, according to MORGENSTIERNE a loanword from an Indo-Aryan language, “Skt *kuñcita-*” (MORGENSTIERNE 1927:33). Sanskrit *kuñcita-* ‘sesame’ (or any plant name), however, does not exist. The Sanskrit *kuñcita-* derives from *kuñc-* ‘to make crooked, to bend or curve’ and means ‘bent, crooked’, which has nothing to do with sesame, whose seeds and plant are neither bent nor particularly curved. The word for ‘sesame’ is *tila-* in Sanskrit.

9.3. The Tocharian forms *kuñcit* and *kwäñcit*

Beyond Iranian, there is for example Tocharian B *kuñcit* ~ *kwäñcit* ~ *kuñcīt* ‘sesame’, Tocharian A *kuñcit* ‘id.’⁴⁹ According to ADAMS (2013:193), both forms derive from Khotanese *kumjsata-* “or [a] similar Iranian source”. Old Uighur and Turkish *künçit* ‘sesame’ was borrowed from Tocharian, but not necessarily from Tocharian A as claimed by CLAUSON (1972:727–728).

Khot. *kumjsata-* can hardly be the source of Tocharian A and B *kuñcit*: *kumjsata* [kundzata] or maybe rather [kundzaʔa] would probably have been rendered in Tocharian A and B as †*kuntsāt*. In any case, Khotanese [dz] would not be rendered in Tocharian as <c> (č). The *ku-* ~ *kwä-* alternation is peculiar in the sense that it is not linked to a change in the placement of the stress of the word. It is quite clear that *kwäñcit* is stressed on the second syllable, and this must also be the case in *kuñcit*. In fact, contrary to our example, most *Cwa-* ~ *Cu-* alternations are linked to the accent being on the first syllable (in *Cwa-*

⁴⁷ As seen below, Iranian variants such as Cl. Pers. *kunjd* indicate, in my opinion, that the <y> in the written form <kwncy> is not there as a *mater lectionis*, but represents a form *kunjt*, with a long *i*.

⁴⁸ In New Persian ‘sesame’ is also called *samsam* (borrowed from Arabic), a form ultimately related to Lat. *sesamum* (the source of the English word), itself from Greek σήσαμον, a borrowing from a Semitic language (cp. Ugaritic *ššmn*, DULAT 847–848).

⁴⁹ Also mentioned in Tocharian are: 1. sesame oil, e.g. IOL Toch 306 a5, *šalype kuñcitäšše* ‘sesame oil’ (see CARLING 2003:76, 91), 2. sesame paste, IOL Toch 306 a6 *kuñcīt mlyokotau* (id., p. 83).

/Cwá/) and then moving to another syllable (/Cu/), cf. FELLNER (2006) for more information. There are only three parallel cases: *kwälyp-* ~ *kulyp* ‘to desire’, an inherited verb (ADAMS 2013:197), *kwäntsaiñne* ~ *kuntsaiñne* ‘firm’ (FELLNER 2006:56) and *kurkam** ~ *kwärkam**.⁵⁰ BAILEY (1937:913) relates the variation of Tocharian B *kuñcit* ~ *kwäñcit* : Ir. *kunčit* to that of TB *twānkaro* : Khot. *ttuṃgara-* ‘ginger’⁵¹ and TB *kurkam** ~ *kwärkam** : Khot. *kurkuma-*, *kurkuṃ*, MP <kwrkwm>, Turkish *kürküm* ‘saffron’.⁵²

According to TREMBLAY (2005:428) TB *twānkaro* was borrowed from an older form **tuṃam-karä*, based on the reconstruction of *ttuṃgara-* by BAILEY (DKS 130). If BAILEY’s (and TREMBLAY’s) etymology is correct, TB *twānkaro* represents a different problem to both *kuñcit* ~ *kwäñcit* and *kurkam** ~ *kwärkam**. First, *twānkaro* does not alternate within Tocharian between *twānkaro* and **tuñkaro*: such an alternation is impossible, the noun being /twānkaro/. Second, *twānkaro* is stressed on the first syllable, while the two other words are stressed on the second syllable.⁵³ Both points seem to indicate that TB *twānkaro* was borrowed at a quite early stage, possibly from Pre-Khotanese, especially if it were borrowed under the form **tuṃam-karai* as suggested by BAILEY (see fn. 51). As stated above, the Tocharian word ‘sesame’ cannot have been borrowed from (any stage of) Khotanese, and I believe the same can be said about the Tocharian word for ‘saffron’: had it been borrowed from Khotanese it would be †*kurkuṃ* or †*kurkuma*.

⁵⁰ FELLNER does not mention *kurkam** ~ *kwärkam**. Furthermore he mentions (p. 56) *kwarsarwa* ~ *kursarwa* (plural forms of *kwarsär* ‘mile, vehicle’), which are in fact *kwärsarwa* ~ *kursarwa* (I thank Dr. Alessandro DEL TOMBA for pointing out this fact to me). The form *kwarsarwa* apparently occurs in IOL Toch 886b4, edited by TAMAI: http://idp.bl.uk/database/oo_cat.a4d?shortref=Tamai_2007. Even if TAMAI’s reading were correct, it must be that, either by mistake or for scribal reasons, the scribe wrote <a> instead of <ä>, since the underlying word is /kwärsärwa/. Stress in both the first and second syllable is, naturally, impossible. The case of *kwärsarwa* is nevertheless different from our other examples, since *kwärsarwa* ~ *kursarwa* represents the plural of *kwarsär*, and does not represent alternative forms in *ku* ~ *kwä* of a single word, originally in *ku-*. Simply, when the word became trisyllabic in the plural, stress shifted to the second syllable. It is thus in fact the *ku-* in the sg. *kursarwa* that needs to be explained.

⁵¹ According to BAILEY (1937:913), the Tocharian form was first borrowed as *twānkarai* (**ttuṃgarai* < **ttuṃgaraka-*), which was reinterpreted as the oblique, after which the nom. *twānkaro* was remade analogically, according to the model nom. *okso*: obl. *oksai* ‘ox’.

⁵² The word *kwärkam** ~ *kurkam** is not attested as such, but its derived adjective *kwärkamäṣṣi*/*kurkamäṣṣi* ‘made of saffron’, ‘saffron-ey’ is. While *kurkamäṣṣi* is attested eleven times in the Tocharian corpus, *kwärkamäṣṣi* is only attested once, in W29 b1.

⁵³ While the stress can only be seen in the forms *kwäñcit* /kwäñcit/ and *kwärkam** /kwärkám/, (and, maybe, in *kuñcit* where the <ñ> could possibly indicate stress), the reduction of *wä* to *u* cannot be due to a shift of stress towards the first syllable, unlike most other TB examples which operate in the opposite direction (stressed *u* > *wa* /wá/).

The word *kuñcit* must then derive from an Iranian language in the region of Tocharian, which is neither Khotanese nor Tumshuqese, and probably also not Sogdian (cf. the Sogdian forms cited above). There is also a possibility of a non-Iranian source for this word, but this seems less probable, given that this word, in antique times, is mostly found in Iranian languages, and it probably found its way into Tocharian thanks to trade with Iranian merchants.

I have no explanation for the alternation *ku-* : *kwä-* in *kwäñcit*, *kwälyp-*, *kwäntsaññe* and *kwärkam**, which is an internal Tocharian problem (see the treatment in FELLNER 2006).

9.4. Etymological discussion

Now, since the etymology of this word is unclear, it could be useful to look for similar words in other languages. The closest word I could find is Sanskrit *kuñci-* ‘Kümmel’ (EWAia³ 95–96) and Sanskrit *kuñcikā* f., which designates two different plants: 1. ‘*Nigella indica* = Schwarzkümmel’ and 2. ‘*Arbus precatorius* = jequirity beans’ (id.). Since jequirity beans are poisonous and red, while cumin is yellow and edible, it seems more probable that the Iranian word *kunčit*, possibly from Old Iranian **kunčita-* has some connection to the first word rather than to the second. Both words could thus be from the Indo-Iranian substratum (as defined by LUBOTSKY 2001), deriving from a root **kunči-*. Nevertheless, cumin and sesame are very different, and the semantic connection is unclear to me. There is variation in the final syllable of this word *-it* ~ *-īt* in almost all Iranian languages (and *-ata* in Khot.). Could this indicate the existence of a variation **kunčīta* ~ **kunčita* in Proto-Iranian perhaps due to the *i* vowel of the donor language being different from those of the Iranian phonological systems? If the word is old (as in at least Proto-Iranian), another solution should be preferred (see Section 6.4.2). Malay and Indonesian *kunyt* ‘turmeric’, which was borrowed as Dutch *koenjit* ‘id.’, (Old) Javanese *kunir*, (probably the source of) Surinamese Dutch *koenier*, all meaning ‘turmeric’, appear related to OIr. **kunči-* on the formal level. The form can be reconstructed for Proto-Malayo-Polynesian, **kunij* ([kunig^j]) ‘turmeric’ (also meaning ‘yellow’ in a number of languages), cf. BLUST & TRUSSELS, ACD. Proto-Malayo-Polynesian predates any Indo-European influence. Nevertheless, ‘turmeric’, ‘sesame’ and ‘cumin’ differ both in appearance and use. Although a formal link between all those forms seems to be phonetically possible, it should not overshadow the fact that a semantic link is very difficult to find.⁵⁴

⁵⁴ I thank Dr. Yvonne VAN AMERONGEN for her help with these words.

10. *totorog* Gav. of Yazd فاخته = ‘turtledove’ (SFB)

The noun *totorog* designates the turtledove, a type of bird related to doves and pigeons. Through internal reconstruction, I posit the form **tutur-* in Proto-Gavruni, with the suffix *-og* that has different functions and is often used in the derivation of plant and animal names.

This word is clearly in its origin an onomatopoeia imitating the song of the turtledove (cf. Ugaritic *tr* ‘turtledove’ DULAT 876, Hebrew תור /to:r/). This word is a mimophonic cognate of Latin *turtur* ‘turtledove, dove’, Polish *turkawka* (< **tur-kaw(-)ka-*), and English *turtle-* in *turtledove* (obviously not related to the turtle). On the basis of all other Indo-European and non-Indo-European forms, I reconstruct (Old Iranian) mimophonic **turtur-* as a predecessor of Proto-Gavruni **tutur-*, which underwent a dissimilation of the *r* in the first syllable.

Conclusion

In this paper I have presented nine Gavruni words, all animal or plant names, and discussed their etymology. Since Gavruni speakers were traditionally agriculturalists and farmers, a considerable amount of Gavruni vocabulary concerns farming and agriculture. This vocabulary has not been presented here, but I believe it deserves a similar study.

None of the nine words discussed in the present study are particularly isolated within Iranian: they all find cognates in Iranian languages or beyond, but most of them represent huge etymological difficulties. I chose these words in Gavruni, because it is one of the Iranian languages I know best – despite its difficulties, but also, and especially, because it is one of the lesser-known Iranian languages that deserve to be looked at more systematically.

I focused on various etymological problems, but especially on the substratum question, inspired by Professor LUBOTSKY’s famous (2001) paper. In particular, *kalpak* ‘lizard’, *karpura* ‘(a type of) germander’ and *kuncit* ‘sesame’ proved particularly interesting, but all the results this research yielded, I fear, are speculations and some new tracks to follow. As a conclusion to both the etymology of *kalpak* (Section 6.5.) and *karpura* (Section 8.3.2.), I proposed the reconstruction of two roots: **kalp* or **karp* ‘disgusting animal’ and **karp* ‘odourous plant’. Such a conclusion is not, by any means, elegant, and I hope that further research will prove this particular point wrong, and either connect or disconnect these two roots in some way.

In Section 2, I rejected an Iranian etymology in Arabic, proposed in previous scholarship, preferring an inner-Arabic origin to an Arabic word, itself the source of a Kermān Gavruni word: *poxtu* ‘turtledove’. The non-borrowed name of the turtledove in Gavruni is onomatopoeic in origin, and discussed in Section 10. I also discussed the – probably onomatopoeic – name of the frog in Gavruni in Section 3, which led me to also discuss the name of the frog in various other Iranian languages. In Section 4, I discussed a derivative of this word, the Gavruni name of the mole-cricket. Section 5 was dedicated to the analysis of a very different type of animal, the rooster, and its etymology.

In the Section 6, while discussing the etymology of Gav. *kalpak*, I tried to discuss further the Indo-Iranian suffix **-ā́ća-* which should be reconstructed as **-ā́ća-* or possibly rather **-ā́ća-*. I try to explain variation in length in substratum words and suffixes in Iranian by the effect of internal *sandhi*. This process is probably very old.

In Section 7, I discussed the name of the spider in Gavruni, and in multiple Iranian languages: it is a compound containing the word *kara* ‘spider web’, a word that is perhaps connected to *kar-* ‘to do’.

All of these etymologies led me to discuss a number of words in multiple languages. My utmost wish is that this paper has served its purpose: to interest the reader in both the riches of little-known Iranian languages, and that of the incredible material vocabulary of Indo-Iranian.

ABBREVIATIONS

- ACD = BLUST, Robert & TRUSSEL, Stephen. *The Austronesian Comparative Dictionary*. Web edition (revised 10 July 2019).
- CPD = MACKENZIE, David Neil. 1986. *A Concise Pahlavi Dictionary*. London: Oxford University Press.
- DKS = BAILEY, Harold W. 1979. *Dictionary of Khotan Saka*. Cambridge: Cambridge University Press.
- DULAT = DEL OLMO LETE, Gregorio & SANMARTÍN, Joaquín. 2003. *A Dictionary of the Ugaritic Language in the Alphabetic Tradition. Part one*. Translated by WATSON, Wilfred G.E. Leiden • Boston: Brill.
- EWAia¹ = MAYRHOFER, Manfred. 1992. *Etymologisches Wörterbuch des Altindoarischen*, Band I. Heidelberg: Carl Winter – Universitätsverlag.
- EWAia³ = MAYRHOFER, Manfred. 2001. *Etymologisches Wörterbuch des Altindoarischen*, Band III. Heidelberg: Carl Winter – Universitätsverlag.
- SFB = see SORUŠYĀN 1978.

EWN = *Etymologisch Woordenboek van het Nederlands*. 2003-2009. Edited by Philippa, Marlies, Debrabandere, Frans, Quak, Arend, Schoonheim, Tanneke and van der Sijs, Nicoline. Amsterdam: Amsterdam University Press.

BIBLIOGRAPHY

- ABAEV, Vasily I. 1956. “Скифский Быт и реформа зороастра” *Archiv Orientalní* XXIV, 23-56.
- ADAMS, Douglas Q. 2013. *A Dictionary of Tocharian B. Revised and Greatly Enlarged*. Amsterdam • New York: Rodopi.
- AMIGUES, Suzanne. 2012. “κάρπιον” In: *Chronique d’étymologie grecque n°12*, edited by Lamberterie, Charles de, *Revue de Philologie, de littérature et d’histoire anciennes* LXXXIII, 306.
- ANDERSON, S.C. 1989. “Amphibians” In: *Encyclopædia Iranica* I:9, 987-990.
- ANONBY, Erik and ASADI, Ashraf. 2014. *Bakhtiari Studies. Phonology, Text, Lexicon*. Uppsala: Acta Universitatis Upsaliensis.
- ASATRIAN, Garnik. 2009. “Iranian notes III” *Iran and the Caucasus* 13, 319–329.
- BAILEY, Harold W. 1937. “Ttaugara”, *Bulletin of the School of Oriental Studies* VIII:4, 883–992.
- 1970. “A range of Iranica” In: *W. B. Henning Memorial Volume*. Edited by Boyce, Mary and Gershevitch, Ilya. London: Lund Humphries, 20–36.
- BAIZOYEV, Azim and Hayward, John. 2004. *A Beginner’s Guide to Tajiki*. London • New-York: Routledge.
- BARTHOLOMAE, Christian. 1904. *Altiranisches Wörterbuch*. Strassburg: Verlag von Karl J. Trübner.
- BEEKES, Robert S. P. 1988. *A Grammar of Gatha Avestan*. Leiden • New-York • København • Köln: Brill.
- BELLAMY, Kate. 2018. *On the external relations of Purepecha. An investigation into classification, contact and patterns of word formation*. Utrecht: LOT.
- BENVENISTE, Émile. 1940. *Textes sogdiens, édités, traduits et commentés*. Paris: Paul Geuthner.
- BENZING, Johannes. 1983. *Chwaresmischer Wortindex*. Edited by Taraf, Zahra. Wiesbaden: Otto Harrassowitz.
- BERNARD, Chams B. 2016. *Le Gavrouni, Enquête phonologique et étude comparée*. Masters thesis. Paris: EPHE.
- BOILEAU DESPRÉAUX, Nicolas. [1674] 1825. *L’Art Poétique de Boileau Despréaux, avec des éclaircissements historiques donnés par lui-même et les remarques de Cl. Brossette*. Compiègne: Imprimerie de Jules Escuyer.

- BOYCE, Mary N. E. 1969. "Some aspects of farming in a Zoroastrian village of Yazd" *Persica* IV, 121–140.
- CARLING, Gerd. 2003. "New look at the Tocharian B medical manuscript IOL Toch 305 (Stein Ch.00316.a2) of the British Library – Oriental and India Office Collections" *Historische Sprachforschung* 116:1, 75–95.
- CHEUNG, Johnny. 2002. *Studies in the Historical Development of the Ossetic Vocalism*. Wiesbaden: Dr. Ludwig Reichert Verlag.
- 2007. *Etymological Dictionary of the Iranian Verb*. Leiden • Boston: Brill.
- CHRISTENSEN, Arthur. 1930. *Contributions à la dialectologie iranienne. Dialecte guiläkī de Recht, dialectes de Färizānd, de Yaran et de Natanz*. København: Det Kgl. Danske Videnskabernes Selskab.
- CLAUSON, Gerard. 1972. *An Etymological Dictionary of Pre-Thirteenth-Century Turkish*. Oxford: Clarendon.
- DE VAAN, Michiel. 2000. "The Indo-Iranian Animal Suffix *-Āćá*", *Indo-Iranian Journal* 43, 279–293.
- 2003. *The Avestan Vowels*. Amsterdam • New York: Rodopi.
- DOERFER, Gerhard. 1975. *Türkische und mongolische Elemente im Neupersischen. Band IV Türkische Elemente im Neupersischen (Schluss) und Register zum Gesamtwerk*. Wiesbaden: Franz Steiner.
- DONKIN, Robin A. 1999. *Dragon's Brain Perfume. An Historical Geography of Camphor*. Leiden • Boston • Köln: Brill.
- EILERS, Wilhelm. 1971. "Iranisches Lehngut im Arabischen" In: *Actas IV Congresso de estudos árabes e islâmicos*. Leiden: Brill, 581–657.
- FARSANI, Mohsen. 2011. *Étude Lexicologique de la langue bakhtiari d'Iran*. PhD thesis. Paris: Sorbonne Nouvelle (Paris 3).
- FELLNER, Hannes. 2006. "On the Development of Labiovelars in Tocharian" In: *Proceedings of the Seventeenth Annual UCLA Indo-European Conference, Los Angeles 2005*. Edited by Jones-Bley, Karlene, Huld, Martin E., Della Volpe, Angela and Robbins Dexter, Miriam.
- FERLUS, Michel. 1977. "L'infixe instrumental *rn* en khamou et sa trace en vietnamien" *Cahiers de Linguistique - Asie Orientale* 2, 51–55.
- GERSHEVITCH, Ilya. 1954. *A Grammar of Manichean Sogdian*. Oxford: Basil Blackwell.
- GHARIB, Badrozzamān. 2004. *Sogdian Dictionary. Sogdian - Persian - English*. Tehran: Farhang Publications.
- GHOLAMI, Saloumeh and FARAHMAND, Armita. 2016. *Zoroastrian Dari (Behdini) in Kerman*. Girona: Estudios Iranios y Turanios.

- GODDARD, Ives. 1974. "An Outline of the Historical Phonology of Arapaho and Atsina" *International Journal of American Linguistics* 40:2, 102–116.
- GOLFAKHRABADI, Fereshteh, YOUSEFBEYK, Fatemeh, MIRNEZAMI, Tahmineh, LAGHAEI, Pedram, HAJIMAHMOODI, Mannan, KHANAVI, Mahnaz. 2015. "Antioxidant and Antiacetylcholinesterase Activity of *Teucrium hyrcanicum*" *Pharmacognosy Res* 7 (Suppl 1), June 2015, S15–S19.
- GONDA, Jan. 1932. *Austrisch en Arisch. Het belang van de kennis der austrische talen, voornamelijk voor de indische philologie*. Utrecht: De Industrie.
- HAMMER, Niels. 2019. "The Evolution of a Species-Specific Term: Old Indic *Krauñcaḥ* (*Grus grus*)" *Zeitschrift der Deutschen Morgenländischen Gesellschaft* 169/2, 395–408.
- HENNING, Walter B. 1933. "Das Verbum des Mittelpersischen der Turfanfragmente" *Zeitschrift für Indologie und Iranistik* 9, 158–253.
- 1939. "Sogdian Loan-words in New Persian" *Bulletin of the School of Oriental Studies* X:1, 93–106.
- 1947. "Two Manichaean Magical Texts with an Excursus on The Parthian ending *-ēndēh*" *Bulletin of the School of Oriental Studies* XII:1, 39–66.
- HOFFMANN, Karl & FORSSMAN, Bernhard. 2004. *Avestische Laut- und Flexionslehre. 2. durchgesehene und erweiterte Auflage*. Innsbruck: IBS.
- HOUTUM-SCHINDLER, Albert. 1882. "Die Parsen in Persien, ihre Sprache und einige ihre Gebräuche" *ZDMG* 36, 54–88.
- IVANOW, Vladimir A. 1939. "The Gabri Dialect spoken by the Zoroastrians of Persia - III" *Rivista degli Studi Orientali* 18/1, 1–57.
- KENT, Roland G. 1953. *Old Persian: Grammar, Texts, Lexicon*. New Haven: American Oriental Society.
- KHADIGE, A., KESHA VARZ Z., MOJAB F. and MAJD, HA. 2017. "The effect of *Teucrium polium* on the duration of menstrual bleeding: A triple-blind placebo-controlled clinical trial", *Electron Physician*, 9(9), 5233–5236.
- KORN, Agnes. 2005. *Towards a Historical Grammar of Balochi*. Wiesbaden: Dr. Ludwig Reichert Verlag.
- LANE, Edward W. 1865. *An Arabic-English Lexicon, Derived from the best and most copious Eastern sources*. London: Williams and Norgate.
- LUBOTSKY, Alexander. 2001. "The Indo-Iranian substratum" In: Chr. Carpelan, A. Parpola, P. Koskikallio (eds.) *Early Contacts between Uralic and Indo-European: Linguistic and Archaeological Considerations. Papers presented at an international symposium held at the Tvärminne Research Station of the University of Helsinki 8-10 January 1999*, Mémoires de la Société Finno-ougrienne, Vol. 242. Helsinki: Société Finno-ougrienne.

- MALANDRA, William. 2010. "Karapan" In: *Encyclopedia Iranica Vol. XV, Fasc. 5*, p. 550. Accessed online on the 14th of November 2019.
- MAZDĀPOUR, Katāyun. 2006. *Vāže-nāme-ye guyeš-e Behdinān-e šahr-e Yazd (vol. II)*. Tehran: Pažuhešgāh-e 'Ulum-e Ensāni va Motale'āt-e Farhangī.
- MORGENSTIERNE, Georg. 1927. *An Etymological Dictionary of Pashto*. Oslo: Hos Jacob Dybwad.
- 1938. *Indo-Iranian Frontier Languages vol. II. Iranian Pamir Languages*. Oslo: Instituttet for sammenlignende kulturforskning.
- 1973. *Irano-Dardica*. Wiesbaden: Dr. Ludwig Reichert Verlag.
- NOVÁK, Lubomír. 2013. *Problem of Archaism and Innovation in the Eastern Iranian Languages*. PhD thesis. Prague: Charles University.
- NYBERG, Henrik S. 1974. *A Manual of Pahlavi II. Ideograms, Glossary, Abbreviations, Index, Grammatical Survey, Corrigenda to Part I*. Wiesbaden: Otto Harrassowitz.
- PETERMANN, Julius Heinrich. 1865. *Reisen im Orient*. (Zweite Ausgabe). Zweiter Band. Leipzig: Verlag von Veit & Comp.
- POU, Saveros. 2004. *Dictionnaire vieux-khmer - français - anglais*. Paris: L'Harmattan. (First edition 1992. Paris: Cedoreck).
- SALEMANN, Carl H. 1895–1901. "Mittelpersisch" In: *Grundriss der Iranischen Philologie. Erster Band. 1. Ableitung*. Edited by Geiger, Wilhelm and Kuhn, Ernst. Strassburg: Verlag von Karl J. Trübner, 249–332.
- SCHAPKA, Ulrich. 1972. *Die Persischen Vogelnamen*. PhD thesis: Würzburg.
- SCHLIMMER, Johann L. 1874. *Terminologie médico-pharmaceutique et anthropologique française-persane*. Tehran.
- SCHWARTZ, Martin. 2013. "Avestan *kauui-* and the Vocabulary of Indo-Iranian Institutions" *Gifts to a Magus, Indo-Iranian Studies Honoring Firoze Kotwal*. Edited by Choksy, Jamsheed and Dubeansky, Jennifer. New-York: Peter Lang Publishing.
- SHORTO, Harry. 2006. *A Mon-Khmer Comparative Dictionary*. Edited by Sidwell, Paul, assisted by Cooper, Doug and Bauer, Christian. Canberra: the Australian National University.
- SORUŠYĀN, Jamšid S. 1978. *Farhang-e Behdinān*. Edited by Sotoudeh, Manuchehr. Tehran: Enteshārāt-e Dānešgāh-e Tehrān.
- SOTOUDEH, Manuchehr. 1956. *Farhang-e Kermāni*. Tehran: Enteshārāt-e Farhang-e Irānzamin.
- STEINGASS, Francis J. 1892. *A Comprehensive Persian-English Dictionary, Including the Arabic words and phrases to be met with in Persian literature, being Johnson and Richardson's Persian, Arabic and English Dictionary*,

revised, enlarged, and entirely reconstructed. London: Routledge and Kegan Paul limited.

TREMBLAY, Xavier. 2005. "Irano-Tocharica et Tocharo-Iranica" *Bulletin of the School of Oriental and African Studies* 68:3, 421-449.

VOSKANIAN, Vardan and BOYAJIAN-SURENIANTS, Vahe. 2007. "Lexical Gleanings from Bašākard" *Iran and the Caucasus* 11, 121-125.

WITZEL, Michael. 1999. "Substrate Languages in Old Indo-Aryan". *Electronic Journal of Vedic Studies* 5-1, 1-67.

XOMAK, Javād-Mohammad. 2000. *Vāže-nāme-ye Sagzi*. Tehran: Soruš.

