Persian *ezāfe* as a contact-induced feature

© 2020   **Ilya S. Yakubovich**

Institute of Oriental Studies, Russian Academy of Sciences,
Moscow, Russia; sogdiana783@gmail.com

Abstract: The paper addresses the origin of the right-branching morphosyntactic pattern known as the *ezāfe* construction, which permeates the Persian noun phrase syntax. In Section 1, I attempt to demonstrate that this construction poses challenges for word order typology. Furthermore, it represents a deviation from the pattern of left-branching, predominant in the nominal syntax of those modern Iranian languages that avoided intensive contacts with Persian. In Section 2, I extend the analysis to Old and Middle Iranian and argue that the right-branching construction with the relative pronoun *haya-*/*taya-* functioning as a linker, the ancestor of the later *ezāfe* construction, was grammaticized in Old Persian. Section 3 constitutes the core of this paper, and presents arguments for the influence of the Elamite language on syntactic restructuring in Old Persian. First, Elamite shares with Persian the unusual combination of the basic SOV word order and right-branching nominal syntax. Second, Elamite constructions with class agreement could plausibly be calqued by noun phrases with the Old Persian relative pronoun *haya-*/*taya-* in the linking function. Third, the demonstrable language shift from Elamite to Iranian along the northern shore of the Persian Gulf in the first millennium BCE provides a suitable sociolinguistic correlate for partial syntactic restructuring in Old Persian. Section 4 tackles the arguments for Urartian, alongside Elamite, affecting Iranian nominal syntax, and pleads for their inconclusive character. Section 5 addresses noun phrases in Bactrian, Khwarezmian, and Younger Avestan, which may potentially be used as evidence against the proposed scenario, and offers their alternative interpretations.

Keywords: Avesta, Bactrian, Elamite, *ezafe*, Farsi, Hurro-Urartian, Iranian, Khwarezmian, language contacts, language shift, Old Persian, word order

Персидский изафет как продукт языковых контактов

Илья Сергеевич Якубович

Институт востоковедения РАН, Москва, Россия; sogdiana783@gmail.com

Аннотация: Статья посвящена происхождению так называемой изафетной конструкции в современном персидском языке, характеризующейся правым ветвлением в именных синтагмах. В первом разделе обсуждается проблематичный характер данной конструкции с точки зрения типологии

порядка слов, а также ее контраст с левоветвящимися именными синтагмами, которые характерны для большинства иранских языков, не подверженных интенсивному влиянию персидского. Во втором разделе, посвященном анализу соответствующих именных словосочетаний в древне- и среднеиранских языках, делается вывод о грамматикализации правоветвящейся конструкции с относительным местоимением *haya-/taya-* в качестве связки, предка современной изафетной конструкции, в древнеперсидском языке. Ядром настоящей статьи является ее третий раздел, обосновывающий гипотезу о синтаксической перестройке древнеперсидского языка под эламским влиянием. Во-первых, эламский язык обнаруживает нетривиальное сочетание базового порядка слов SOV и правого ветвления в именной системе, характерное также и для древнеперсидского. Во-вторых, использование относительного местоимения *haya-/taya-* в функции связки в древнеперсидском языке может являться калькой эламского механизма классного согласования. В-третьих, несомненный переход носителей эламского языка на иранские языки в регионе к северу от Персидского залива в первом тысячелетии до н. э. обеспечивает социолингвистические предпосылки для контактно-обусловленных синтаксических изменений в древнеперсидском. В четвертом разделе обсуждается вопрос о возможном влиянии урартского языка, наряду с эламским, на эволюцию иранского именного синтаксиса и делается вывод о бездоказательном характере данной гипотезы. В пятом разделе разбираются примеры бактрийских, хорезмийских и младоавестийских именных словосочетаний, которые могут служить потенциальными аргументами против предложенного сценария, и предлагаются их альтернативные интерпретации.

1. Noun phrase in New Iranian

A feature of the Persian language which remains stable from the classical period onwards is the right-branching nominal syntax. The dependent nouns recursively follow their syntactic heads, while the adjectives precede the nouns depending on the same head. A special linker, known as the *ezāfe* in Persian grammatical tradition, is placed between the head and the dependent and cliticized to the head.[1] In the absence of the grammatical case category in Persian, this device represents the primary means of ensuring cohesion within a noun phrase. Examples (1)–(3) are taken from the Classical Persian poets Rudaki, Ferdowsi and Hafez respectively, but neither the poetic character nor the medieval origin of the respective compositions impact the nature of the *ezāfe* construction. In modern prose or colloquial language, it remains essentially the same, except that the shape of the *ezāfe* particle is *-e* rather than *-i* in contemporary Persian (Farsi) and Dari of Afghanistan, while in Tajiki it stays as *-i*. There is a scholarly consensus that this particle goes back to the Proto-Iranian relative pronoun [Haider, Zwanziger 1984, with further references].

(1)   yād=i   yār=i   mihrbān   āy-ad   hamē
        memory=LNK   friend=LNK   dear   come.PRS-3SG   IMPF
        'The memory of a dear friend is coming (back)'.

(2)   ba   nām=i   xʷadāwand=i   jān   u   xirad
        in   name=LNK   lord=LNK   soul   and   reason
        'In the name of the lord of soul and reason…'

---

[1] A more detailed synchronic discussion of the Persian *ezāfe* construction, accompanied by numerous examples, can be found in [Rubinchik 2001: 365–368]. Note that the Arabic *iḍāfa* and the Turkish *izafet* have nothing in common with the Persian *ezāfe* except the name (which is Arabic in origin). These are three different ways of organizing nominal syntax in the respective languages.

(3) bāzār=i   but-ān   šikast gīr-ad
     market=LNK   idol-PL   defeat receive.PRS-3SG
     'The fair of idols suffers defeat.'

As one can infer from examples (1) and (3), the right-branching *ezāfe* constructions coexist with the basic SOV word order in Persian. Nevertheless, it represents a typologically non-trivial feature, as noted already in [Greenberg 1963: 67], where Persian is listed under the rare Type 17. The seminal work of Matthew S. Dryer on word order in the world's languages contains a sample of 124 languages characterized by the OV word order, out of which 112 show the preposition of the nominal possessor to the head noun, while only 12 feature the possessor in the postposition [Dryer 1992: 91]. The statistics of [Dryer, Haspelmath 2013] lists 26 languages characterized by the correlation Noun-Genitive / SOV, as opposed to 249 languages featuring Noun-Genitive / SVO, 106 languages featuring Genitive-Noun / SVO, and 398 languages featuring Genitive-Noun / SOV (https://wals.info/combinations/86A_81A#2/24.9/153.0).

Within the Iranian family, Persian is not the only language with the *ezāfe* construction as the principal way of organizing nominal syntax. It is prevalent in the West Iranian languages Kurdish and Zazaki, where the linker is notoriously inflected [Efimov 1975: 78–82]. It is also fundamental to Dari and Tajiki, the erstwhile varieties of Persian whose development into independent languages was precipitated by sociolinguistic factors, as well as to several vernaculars of Iran, which lack written tradition and have been developing under the influence of Persian in the last two centuries. Based on the analysis of this set one can conclude that the *ezāfe* construction was at home on the western margin of the Iranian-speaking area, at least before the spread of Persian to Central Asia in the Early Islamic period complicated the picture.[2]

In contrast, the majority of the modern Iranian languages with limited restructuring under Persian influence display the typologically common pattern of SOV word order combined with the preposition of the dependent noun in noun phrases ("Genitive-Noun"). Such a state of affairs can be illustrated with reference to the West Iranian language Balochi (4), as well as the East Iranian languages Wakhi (5) and Ossetic (6). This genetic diversity correlates with a wide geographic spread: the Balochi language is spoken at the border of Afghanistan and Pakistan, the Wakhi language is at home in the Pamir Mountains, while the Ossetians inhabit parts of the Caucasus. The selection of examples here is admittedly subjective; for a more comprehensive list of languages that display the same correlations one should consult [Stilo 2005: 47–51].

(4) man wat-ī   laṭṭ   u   kawš-anī   badal-ā   gīr-īn
     I   self.GEN stick and shoe.GEN.PL   exchange.OBL take.PRS-1SG
     'I will take revenge for being beaten up', lit "for stick and shoe" [Jahani, Korn 2009: 671].

---

[2] The so called "inverted *ezāfe*" construction, which exists in certain Northwest Iranian languages of the Caspian area, was recently treated in [Ivanov, Dodykhudoeva 2017]. Its typical examples are Mazanderani noun phrases *Hosayn-e ketāb* 'Husein's book', *serx-e jeme* 'red shirt'. Unlike its Persian counterpart, this construction is left-branching, for example, the adjective *serx* 'red' precedes the noun *jeme* 'shirt'. One can agree with the authors of this paper that the "inverted *ezāfe*" is typologically more similar to a linker than case marker, in particular because it can be deployed in both possessive constructions and adjectival phrases. In historical terms, however, there is every reason to accept the traditional analysis of Mazanderani *-e* as the reflex of the Iranian genitive case in *-ahya*, which underwent functional extension on the model of the *ezāfe* particle as a result of syntactic interference with Persian. Since the native speakers of Mazanderani and other Caspian languages are usually bilingual in Persian, this convergence is sociolinguistically motivated. On the other hand, as argued in the rest of this paper, the antecedents of the *ezāfe* construction show the right-branching word order from the Achaemenid period onwards, and the reversal of this pattern in the Caspian languages would lack any motivation. Alternatively, as suggested by one of the reviewers, the "inverted *ezāfe*" after the attributive adjectives may simply continue an old suffix (cf. the suffix *-ēn* marking the attributive adjectives in West Balochi). Therefore, this construction is tangential to the main argument of this paper and will not be treated here in any more detail.

(5)   am   ya   potšo   kənd   am   yaw   yaš   δыvу-ətk
       also   the   king   wife   also   he   horse   steal-PERF
       '(He) has stolen the king's wife as well as his horse' [Steblin-Kamenskii 1987: 472].

(6)   woræzmæg   k'æʒæx-i   sær-i   jew   suvællon-i   fæd-tæ   fæwwid-t-a
       Vorazmag   rock-GEN   top-LOC   one   child-GEN   footstep-PL   see-PRT-3SG
       'On top of the rock, Worazmag saw footsteps of a child' (Digoron dialect, [Isaev 1987: 633]).

So far we have seen that the *ezāfe* construction of the Persian type, on the one hand, is typologically unusual, and, on the other hand, was historically confined to a particular area of the Iranian spread zone. Therefore, one can formulate a preliminary hypothesis that it represents an innovation, which has to be tested through the analysis of additional data.

2. Historical perspective

So far we have confined ourselves to the discussion of the New Iranian languages, spoken from the time of the Islamic conquests onwards. The distribution becomes clearer once one takes into consideration the five best attested Middle Iranian languages, used in writing in the first millennium CE. Only one of them, namely Middle Persian, displays the right-branching *ezāfe* construction (linker *ī* or perhaps *ē*) as the most common way of organizing nominal syntax (7). Nevertheless, even in this language it coexists with the alternative strategy of left-branching, which does not require any linker, and is essentially the same as illustrated by New Iranian examples (4)–(6). Thus, the phrase *wāxš ī hwarāsān wimand* 'spirit of the Khorasan frontier' ([Boyce 1975: 40], M2.I §4.1) embeds the dependent-head constituent *hwarāsān wimand* 'frontier of Khorasan' into the *ezāfe* construction. For various types of noun phrases where the *ezāfe* construction was in use in Middle Persian, see [Boyce 1964: 37–47] and [Durkin-Meisterernst 2014: 268–271].

(7)   dar   ī   čašm-ān   kē   pad   dīdišn   īg   tuhīg   wīfs-ēd
       door   LNK   eye-OBL.PL   REL   at   vision   LNK   empty   be.deceived.PRS-3SG
       'Door of eyes (= sight), which is deceived at an empty vision (= mirage)' (M 2 I §7.1, [Boyce 1975: 41]).

The other four Middle Iranian languages, namely Parthian (8), Sogdian (9), Bactrian (10), and Khotanese (11), all select left-branching without a linker as the main strategy of forming noun phrases. This does not exclude the existence of alternative patterns, such as right-branching with the linker *čē* in Parthian Manichaean texts, but "its absence is commoner than its use" [Boyce 1964: 32].[3]

(8)   byd   š'bwhr   š'h-'n   š'h̲   br'd   bwd   (m)yšwn   xwd'y
       furthermore   Shapur   king-s   king   brother   be.3SG.PRT   Mesene   lord
       'Furthermore Shapur, king of kings, had a brother, lord of Mesene' (M 47.I 6–7, [Sundermann 1981: 102]).

(9)   rty   MN   xmyr   w'nkw   pwstkw   ''βr-nt
       furthermore   from   amir   such.ACC   letter   bring-3PL.PRT

---

[3] For the discussion of Classical Armenian relative clauses without overt predicates and their possible areal connections, see [Meyer 2017: 208]. Whether the Parthian influence was responsible in whole or in part for the existence of such constructions in Classical Armenian remains unclear, but even if it did, the elliptical relative clauses introduced by the Parthian particle *kē* represent a more likely source of inspiration than the *ezāfe* construction with the particle *čē*.

pr	xmyr	xws'nty-'kH	c'-β'k
with	amir	satisfaction	from-you

'And they brought from emir this letter with emir's satisfaction with you' (Mugh 1.I 4–5, [Yakubovich 2002: 234]).

(10)
πιδοοασατο	χοακαμο	χοασινδο	βαγοφαρνο	ζαμωδο	π̅ο̅ρ̅ο̅
declare.3SG.PRT	own.wish	own.approval	Bagfarn	Zamod	son

'And Bag-farn, son of Zamod declared freely and willingly…' [Sims-Williams 2000: 33], A 8.

(11)
tt-ye	uysnor-i	käḍägānīne-i	haṃbīs-ä
that-GEN.SG.M	being-GEN.SG.M	karma.related-NOM.SG.M	heap-NOM.SG.M

'The karmic heap of this being…' (SghS. §35.2, [Canevascini 1993: 16]).

It is relevant for this discussion that Parthian is a Northwest Iranian language, whereas Bactrian, Sogdian and Khotanese belong to the East Iranian group. The Parthian language was presumably at home to the southeast of the Caspian Sea, at the border of the present-day Iran and Turkmenistan; Bactrian was spoken on the territory of Northern Afghanistan, Sogdian in Tajikistan and Uzbekistan, and Khotanese in Xinjiang, China. This genetic and geographic diversity confirms the impression that the dependent-head word order was the most common one in the Iranian-speaking area and is to be reconstructed as prototypical for Proto-Iranian nominal syntax.

It is useful to explore this hypothesis further by considering the data of the Old Iranian languages. In Old Persian, attested through the Achaemenid royal inscriptions of the 6th–4th centuries BCE, there are two principal ways of organizing nominal syntax. These are either left-branching phrases, which do not require any linker, as in the predicate group of (12), or their right-branching counterparts, where the syntactic head and dependent are linked by the pronoun *haya-*/*taya-*, which agrees in gender, number, and case with the head, as in (13). In historical terms, *haya-*/*taya-* represents a combination of the demonstrative pronoun **ha-*/**ta-* and relative pronoun **ya-*, but synchronically it is also deployed as the head of normal relative clauses [Skjærvø 2009: 154].[4] The right-branching constructions with the linker display a wide variety of syntactic subtypes, e.g. XPf. 30 *dārayavauš haya manā pitā* 'Darius, my father', DB. I 88–89 *avam kāram tayam Nadintabairahyā* 'that army of Nadintu-Bel', DB. I 80–81 *xšaçam taya Babirau* 'the kingship in Babylon'.

(12)
aita	xšaça-m	hacā	paruviya-ta	amāxam	taumā-yā	āh-a
this.NOM.SG.N	kingdom-NOM.SG	from	beginning-ABL	we.GEN	family-GEN	be.PRT-3SG

'This kingdom from the beginning belonged to our family' (DB. I 45, [Schmitt 2009: 43]).

(13)
kāra	[haya	ma]nā	ava-m	kāram
host.NOM.SG	REL.NOM.SG.M	I.GEN	that-ACC.SG.M	host.ACC.SG

taya-m	hamiçiya-m	a-ja	vasai
REL-ACC.SG.M	rebellious-ACC.SG.M.	PRT-smite.3SG	at_will

'My army smote the rebellious army at will' (DB. II 25–26 [Schmitt 2009: 54]).

There is a tendency to deploy the variant with the linker to mark the pragmatically familiar entities. Thus 'our family' in (12) constitutes new information: Darius, the author of the Bisitun

---

[4] The pronoun *haya-*/*taya-* is formally different from **ya-*, which is traditionally reconstructed as the ancestor of the *ezāfe* particle in Middle and New Persian. This discrepancy is usually understood as a testimony to the possibility of unique innovations in the language of Achaemenid elites. For justifying the rise of the morpheme cluster *ha-ya-* / *ta-ya-* in functional terms, see [Adiego 2000]. The recent attempt to derive the relative pronoun and linker *haya-* / *taya-* via reanalysis of the demonstrative pronoun **ha-*/*ta-* [Brust 2018: 194] does not take into consideration the syntactic properties of *haya-* / *taya-*, which set it apart from a prototypical article or demonstrative pronoun (cf. the discussion above). On the other hand, Nicholas Sims-Williams (pers. comm.) draws my attention to a possibility of deriving the Middle Persian *ezāfe* particle directly from *haya-*, with the irregular loss of *h-* in a clitic formation. He specifies that the derivation of this morpheme from *haya-* vs. **ya-* should be regarded as an open question.

inscription, makes a significant (and probably false) claim that he represents a legitimate heir to the throne. On the contrary, 'my army' and 'rebellious army' in (13) represent information established in the immediately preceding discourse.[5] Such a quasi-anaphoric use is perhaps why the element *haya-*/*taya-* is sometimes called "article" in secondary literature on Old Persian, even though such a designation is ultimately misleading.[6] The construction *kāra haya manā* is commonly regarded as the antecedent of the *ezāfe* construction in later Persian. It is restricted to the nominative, accusative and ablative-instrumental cases and never occurs in genitive-dative, cf. [Probert 2015: 412, with further references].

The basic Old Persian word order at the clause level is SOV [Kent 1953: 96; Skjærvø 2009: 94]. In this sense, it is no different from modern Farsi, or for that matter, from the majority of the other Iranian languages. It is more difficult to generalize about the basic syntax of the Old Persian noun phrase. Both patterns head-dependent and dependent-head are abundantly represented there, with a vague pragmatic distribution (see above). Since the first pattern correlated with the use of an extra pronoun, it can be considered as formally marked. In essence, the syntax of the Achaemenid inscription can be compared with that of Attic Greek prose, where the right-branching construction with the repetition of the article ἡ ἐπιθυμία ἡ τοῦ ὕδατος 'the desire for water' constitutes a marked alternative to the synonymous left-branching phrase ἡ τοῦ ὕδατος ἐπιθυμία. If so, the basic syntax of Old Persian reflects the typologically trivial pattern SOV / Genitive-Noun, which also happens to characterize the majority of the Iranian languages. While the ancestor of the *ezāfe* construction represented a periphrastic alternative to the normal state of affairs, it became approximately as frequent as the inherited left-branching noun phrase in Old Persian.

The other directly attested Old Iranian language is Avestan. It seems appropriate to focus our analysis on the Old Avestan language, because the subsequent stages in the development of the Avesta are increasingly suspect of interference from the Old/Middle Persian vernacular (cf. Section 5). On the other hand, the poetic character of the Old Avestan compositions imposes hurdles on our ability to discern their basic syntax. Nevertheless, it is said that the most common pattern at the clause level is SOV [West 2011: 116] and, at least in the *Yasna Haptaŋhāiti*, the dependent genitives tend to precede their head nouns [West 2011: 123)].

Nevertheless, the sequence of elements that would eventually develop into the *ezāfe* in Persian is already present in Old Avestan, albeit as an exceptional pattern. The Avestan relative pronoun *ya-*, cognate with the second component of Old Persian *ha-ya-*/*ta-ya-*, is found on rare occasions as a noun phrase linker in constructions with right-branching. The difference between such noun phrases and verbless relative clauses consists in the fact that the relative pronouns in these constructions agree in case with their nominal antecedents. The function of the relative in (14) and (15) was synchronically compared with the linking function of the Classical Greek article [West 2011: 82]. Historically, however, there is little doubt that these Old Avestan examples reflect secondary case attraction in verbless relative clauses [Probert 2015: 410-412].

---

[5] The distinction between the ways of marking definite vs. indefinite noun phrases apparently lingers on in the history of Persian. Thus, Early Judaeo-Persian, representing an intermediate stage between Middle and Classical Persian, displays tendencies toward the word order Adjective-Noun in indefinite noun phrases [Paul 2013: 155]. According to [Brunner 1977: 18], a minimal pair in the Middle Persian legal treatise *Hazār dādestān* ("Thousand Judgments") suggests a contrast between the left-branching order in the indefinite noun phrase 'Farrox's sons and descendants' (any of them) and the right-branching order in the definite noun phrase 'the sons and descendants of Farrox' (those mentioned here).

[6] As already observed in [Kent 1953: 85], there are several cases where *haya-* / *taya-* is placed in front of the noun phrase, and the dependent precedes its head. For example, the phrase 'our family' is also attested as *haya amāxam taumā* (DB. I 8). Furthermore, there are instances where *haya-* / *taya-* determines a single constituent e.g. *taya rastam* '(what is) right' (DNa. 7). Notwithstanding this, *haya-* / *taya-* is hardly a prototypical article: it normally cannot be combined with free-standing nouns in the same grammatical case, while its combinations with free-standing adjectives can be regarded as instances of ellipsis.

(14) mąθrəm   yim   hauruuatātō   ašahiiā   amərətātas-cā
     spell.ACC.SG REL.ACC.SG.M integrity.GEN.SG truth.GEN.SG immortality.GEN.SG-and

'a spell, which [is one] of integrity, truth, and immortality' (Y. 31.6, [West 2011: 147]).

(15) tāiš   śiiaoθənāiš   yāiš   vahištāiš
     this.INSTR.PL deed.INSTR.PL   REL.INSTR.PL best.INSTR.PL

'with those deeds that [are] the best' (Y. 35.4, [West 2011: 172]).

The difference between the situations in Old Avestan and Old Persian is that the construction with the relative pronoun functioning as a linker does not yet exhibit a stable pattern in the former case. Examples (14) and (15) lack precise parallels in the rest of the corpus and can be characterized as figures of speech. Neither are they parallel to one another: the demonstrative pronoun is present in (15) but not in (14). The majority of Old Avestan noun phrases with the postposition of the dependent element do not feature any linker [West 2011: 123], while the most common use of *ya-* outside the relative clauses is that of a determiner with no implied linking function [West 2011: 82].

This is the expected state of affairs, since the Old Avestan language was still very close to the reconstructed Common Iranian, while the innovative *ezāfe* construction otherwise characterizes a particular area of the Iranian spread zone. The frequent use of the relative pronoun as a linker can first be discerned in Old Persian, while the relevant construction acquires the default status in Middle Persian and becomes near-obligatory in the contemporary Persian language. It remains to be seen to what extent this process of gradual change is compatible with the influence of external factors.

3. Elamite influence

The idea that the development of the Persian *ezāfe* construction may be contact-induced is in itself not new. Thus, the introduction to the latest influential compendium of Iranian linguistics contains the following observations: "The two documented non-Iranian languages which the earliest immigrating Persian and Median tribes encountered in the west sometime before the 8th century BCE are Urartian at the northern end of the "Zagros" tier, and Elamite at its southern end. Both Urartian and Elamite were right-branching languages. Therefore the Elamite typology has been recognized as the likely source of the innovative Old Persian relative construction *N haya X* 'N who (is) X', which in turn is the source for the *ezāfe* in the Perside languages. Similarly, the likely source at the northern end of the "Zagros" tier was the equally right-branching Urartian typology" [Windfuhr 2009: 28]. Windfuhr's assumption is also followed in a recent Berkeley Ph.D. thesis devoted to the genealogical classification of the Iranian languages [Cathcart 2015: 43].

While areal influence at the margin of the Iranian spread zone is plausible, nobody would surely doubt the possibility of the formation of a new noun phrase pattern due to internal factors. The mere convergence between two right-branching constructions is not statistically significant in itself, because right-branching characterizes about fifty percent of the world's languages. What strengthens the suspicion of external influence is the fact that the development of the construction *N haya X* eventually yielded a typologically unusual syntactic pattern (see Section 1 above). The basic intuition here is that "languages that do not meet the expectations of implicational universals are often found between two languages, groups of languages, or language areas that are more or less opposite to each other" [Stilo 2005: 38]. Thus, in purely synchronic terms, one can claim that modern Persian represents a transitional case between the dialects of Arabic, with their right-branching, and the languages of Central and South Asia, with their left-branching. The same point is stressed in [Haig 2015: 409] with reference to Kurdish and North-Eastern Neo-Aramaic, exhibiting similar patterns of syntactic discrepancy.

From the historical perspective, this synchronic picture is again no more than a first approximation. We know that Sumerian and Elamite, language isolates of the Ancient Near East, displayed the same combination of SOV and the head-dependent order in the noun phrase already before the Iranian migrations to the present-day Western Iran in the early first millennium BCE. The hypothesis that they likewise represented transitional systems at the time is plausible but not provable: while we know that the West Semitic languages were predominantly right-branching already in the second millennium BCE, the languages that were then spoken on the Iranian plateau are simply unknown to us. But the situation in Ancient Mesopotamia offers a suggestive typological parallel to the syntactic change addressed in this paper: Akkadian, a Semitic language that coexisted there with Sumerian in the third and second millennia BCE, features the same typologically marked combination SOV / Noun-Genitive. Despite some voices of dissent, most scholars agree with the claim that the Akkadian word order represents an innovation vis-à-vis the common Semitic pattern VSO ~ SVO / Noun-Genitive, which was due to the Sumerian substrate [Zólyomi 2011: 402, with further references]. Indeed, this example is cited as one of the paradigmatic cases of syntactic change caused by language contact [Thomason 2001: 88].[7]

Therefore, for the purpose of the present discussion, I shall replace Stilo's maxim with the following claim: "languages that do not meet the expectations of implicational universals are suspect of having undergone partial restructuring as a result of substrate or adstrate interference". This broad formulation obviously encompasses situations where both the trigger and target of restructuring are assumed to have met the expectations of implicational universals, as well as the persistence of mixed types in a specific area. In particular, we shall explore the scenario according to which Elamite, itself a language of the mixed syntactic type, affected the dialects on the southwestern periphery of the Iranian spread zone, which were originally left-branching, just as the rest of the Iranian languages. In order to substantiate this hypothesis it is necessary to demonstrate the presence of necessary sociolinguistic conditions for structural contact as well as non-trivial similarity between the noun phrases that supposedly triggered and underwent restructuring.

It is customary to distinguish between two basic modes of language contact involving the situations of language maintenance and language shift. The first situation primarily correlates with the transfer of elements (most commonly, lexical borrowings), the second one is primarily compatible with structural interference. To be sure, the correlation is not quite strict: there are arrested language shifts, which involve the last-minute re-borrowing from the disappearing languages (e.g. the dialects of the "pygmies" in West Africa), and the instances of asymmetrical bilingualism, where the effects on the minority language or basilect are reminiscent of the outcome of language shift (e.g. Asia Minor Greek). But in the majority of cases, partial restructuring of Language A (or a dialect of A) under the influence of Language B is indicative of the shift from B to A (e.g. from Gaelic to English in Ireland). The robust character of restructuring frequently correlates with the short phase of bilingualism and the imperfect learning of A. These theoretical facts are summarized in Table 1, adapted after [Ross 1991].

The next task is to apply the framework outlined above to the coexistence between Iranian and Elamite, as recently outlined in [Tavernier 2018]. The Elamite language was used in writing on a territory stretching from the eastern extension of the Mesopotamian alluvial plain (the area of Susa, present-day Khuzestan) to the highlands of Fars. In the second millennium BCE, there is no evidence for the presence of Iranian speakers in this area, while Elamite competes with Akkadian and Sumerian in the written sphere, in particular, in the regions situated

---

[7] Another case of substrate interference in the ancient Near East yielding a typologically rare syntactic pattern is arguably that of Hattian, a non-affiliated language that was spoken in Central Anatolia in the early second millennium BCE. According to the analysis of [Goedegebuure 2008], Hattian is prototypically a right-branching language, but with numerous deviations from the ideal patterns. Goedegebuure explains this state of affairs by the interference of the Indo-European Anatolian languages of the early second millennium BCE, which are commonly reconstructed as left-branching.

Language contact typology

	Borrowing		Restructuring	
Dominant lang. of bilinguals	Recipient language (RL)		**Source language (SL)**	
Sociolinguistic situation	I (Language Maintenance)	II (Language Shift)	**I (Language Shift)**	II (Language Maintenance)
Domains of transfer	Words, Morphemes	Words	**Phonology, morphosyntax**	Morphosyntax
Examples	Norman French > Middle English	"Pigmy" > Aka	**Gaelic > Irish English**	Turkish > Asia Minor Greek

closer to Mesopotamia. The first Iranian personal names appear in the Elamite sources in the seventh century BCE and become more numerous in the early sixth century. The relevant documents, however, are all composed in Elamite, and the local rulers mentioned in the texts of this period bear Elamite names. This underscores the social dominance of Elamite in the period under discussion.

The situation changed after the formation of the multinational Achaemenid Empire (late 6th century BCE), where Elamite was used in administration alongside several other languages, notably Aramaic and Akkadian). Two administrative centres of the Achaemenid Empire, namely Susa and Persepolis, were situated on the former territory of Elam, but its rulers, at least beginning with Darius I, had Iranian names and were presumably native speakers of Old Persian. This is also the period when the Old Persian language is put in writing: with one exception [Stolper, Tavernier 2007], its use is confined to royal inscriptions, many of which also have Elamite versions.[8] Interestingly enough, the senior scribes mentioned in the Elamite texts of this period have either Iranian or Semitic names, which suggests that Elamite literacy was maintained as homage to tradition, while Elamite ethnicity was no longer associated with high social status. The collapse of the Achaemenid Empire put an end to writing Elamite, although Elamite names also appear in later sources [Stolper 2004: 64]. The main language that is now spoken on the former territory of Elam is Persian (Farsi).

It follows that firstly, there was indeed a language shift from Elamite to Persian and, secondly, this was a gradual process that involved prolonged bilingualism and acculturation (cf. [Henkelman 2011: 584]). Under such conditions, one indeed expects to find some effects of structural interference in Old Persian and perhaps on the subsequent stages of the Persian language, as per the scenario in Table 1 underscored by the bold script. These changes, however, were unlikely to lead to radical restructuring, of the kind typically observed in creole languages, but rather to the modification of certain structural features, just as is commonly assumed in the instance of prehistoric substrate-driven changes in Akkadian. One of such features could indeed be the grammaticalization of the Old Persian pronoun *ha-ya-*/*ta-ya-* as linker in the right-branching noun phrases. Crucially, the proposed scenario does not predict significant lexical interference

---

[8] The Bisitun monumental inscription of Darius I was composed in three languages, Elamite (2 versions), Babylonian, and Old Persian, but the spatial arrangement of this monument leaves no doubt that the first Elamite version of the inscription, not the Old Persian version, was its starting point. At the end of the Old Persian text one finds the claim of Darius I that he commissioned it in the "Aryan"/ "Iranian" script, which had not been done before. While it is not the earliest case in human history when a new writing system was devised as a nationalistic alternative to the existing ones (Anatolian Hieroglyphs and Ugaritic Cuneiform arguably blaze the trail), it is probably the earliest explicit mention of such a practice in written records [Thomason 2001: 6].

from Elamite to Persian, nor is such empirically observed. Therefore, the actual domain of language contact is well compatible with the independent conclusion that Old Persian constituted the target of language shift (cf. [Thomason 2001: 76]).

Turning to the empirical side of the problem, Elamite, like Persian, is an SOV language, while the Elamite noun phrase is characterized by right-branching [Stolper 2004: 84]. At a stage before the beginning of contacts between Elamite and Iranian, exemplified by Middle Elamite (late second millennium BCE), the Elamite language had no structural cases, except in the pronominal system [Stolper 2004: 74]. Class agreement represents the normative way of marking syntactic relationship within a Middle Elamite noun phrase, a pattern that is frequently discussed in connection with the Bantu languages. The Elamite language distinguishes between the locutive (first person), allocutive (second person), animate singular, animate plural, and inanimate noun classes [Stolper 2004: 73].

The paradigmatic alternation between class markers expresses contrast between both inflectional and derivational forms, e.g. *sunki-k* 'I, king' / *sunki-r* 'king' / *sunki-p* 'kings' / *sunki-me* 'kingship'. Their syntagmatic function is arguably even more important, since the co-occurrence of class markers with identical functions ensures cohesion in a noun phrase, as illustrated by the examples below. In (16), the animate class marker *-ri* is attached to the first-person pronoun, indicating its syntactic dependence (otherwise *u-* 'I' appears without a class marker but is capable of inducing the 1sg. agreement marker *-k*). In (17), an inanimate noun triggers an agreement marker *-me* on its syntactic dependents. Finally, in (18), one observes a more complex pattern of agreement: the pronoun *u-* carries both the animate marker *-ri* and the inanimate marker *-me*, referring back to 'god' and 'temple' respectively.

(16) ᵈInšušinak   nap-ir   u-ri
      Inshushinkak   god-ANIM   I-ANIM
      'Inshushinak, my god' [Grillot-Susini 2008: 50]

(17) siyan   ᵈInšušinak-me   husa-me
      temple   Inshushinak-INAN   wood-INAN
      'wooden temple of Inshushinak' [Grillot-Susini 2008: 43]

(18) siyan   ᵈInšušinak   nap-ir   u-ri-me
      temple   Inshushinak   god-ANIM   I-ANIM-INAN
      'temple of Inshushinak, my god' [Grillot-Susini 2008: 46]

The contention of this paper is that the Old Persian construction **baga haya manā* 'my god' (cf. (13) above), represents a calque of Middle Elamite *nap(-ir) u-ri* 'my god'. It shares both the postposition of the syntactic dependent and the use of agreement marker with its Elamite counterpart. Furthermore, the Elamite construction with class agreement can be deployed for possessive, attributive, and appositional phrases [Stolper 2004: 74, 86], and the same is the basic scope of the Old Persian construction with *haya-/taya-* [Kent 1953: 85].[9] The parallel is, of course, not fully precise: Old Persian agreement classes (three genders and two numbers) are different from those in Elamite, the relative pronoun *haya-/taya-* is placed between the head and the dependent, and most crucially, the recursive use of constructions with *haya-/taya-* does not seem to be attested in the published Old Persian texts. But in the absence of an abrupt language shift, the target language tends to select the constructions that are already licensed by its structure for the implementation of the contact-induced changes. We have seen that the relative

---

[9] Cf. the concise treatment of the Old Persian relative pronouns in the linking function in [Skjærvø 2009: 100–101]. Note that [Kent 1953], still representing one of the most authoritative sources on Old Persian syntax, uses the transcription *hya-* / *tya-* for what is now more commonly transcribed as *haya-* / *taya-*. Kent's transcription is still occasionally used in modern scholarship, e.g. [Viti 2015: 320, fn. 203]. For research history and a compromise solution, implying syncope within the history of Old Persian, see [Tavernier 1999].

pronoun *ya-* was occasionally deployed as a linker in right-branching noun phrases in Old Avestan (14)–(15). What functioned as an exceptional device in the language of Zarathustra, and arguably in Proto-Iranian, must have gained in frequency in Old Persian under the impact of the Elamite substrate.

An additional complication of the proposed hypothesis, which is perhaps responsible for the apparent lack of its explicit formulation in earlier literature, consists in further restructuring of noun phrase morphosyntax in Achaemenid Elamite. The equivalent of *kāra* [*haya ma*] *nā* 'my army' in the Elamite version of the Bisitun inscription is ¹*taššu-ip appa* ¹*uni-na* 'troops that (are) mine' [cf. Bae 2001: 123]. This is the word-by-word translation of the Persian turn of phrase: *appa* is the Elamite relative pronoun, while *-na* is the innovative genitive case ending incorporating the historical relative particle *-a* [Stolper 2004: 74, 76]. In order to make sense of this construction and its numerous parallels in Achaemenid Elamite texts, one has to remember that the bulk of Achaemenid Elamite scribes, judging by their personal names, were Iranians, not Elamites. Therefore, the restructuring of Achaemenid Elamite vis-à-vis the earlier stages of Elamite must reflect, above all, imperfect second language acquisition [Henkelman 2011: 586–595].[10] Even if some native Elamite scribes rubbed shoulders with their Iranian colleagues in the royal chancery, they probably had to learn the conventions of the scribal jargon. This is precisely a sociolinguistic setting where one can expect calques from the lingua franca of the scribes, which was presumably Old Persian or something very close. Special cases of reverse contact-induced developments in scribal milieus require a separate research paradigm and have to be studied apart from instances of language shift in oral communication, which is exemplified here by Elamite influence on Old Persian.[11]

Although Elamite, under the proposed interpretation, represented the primary force in the restructuring of the Old Persian nominal system, the contacts between Persian and its western neighbors obviously did not stop at this point. We know that Aramaic dialects were spoken in Mesopotamia in the Parthian and Sasanian periods (2nd century BCE through 7th century CE), while the Arabic language firmly established itself in Iraq and became politically and culturally dominant in Iran with the formation of the Arab Caliphate. In both Aramaic and Arabic, the possessors and attributes normally follow their syntactic heads, and it is very likely that Persian-Aramaic and Persian-Arabic bilingualism contributed to the proliferation of the right-branching noun phrases in Persian (cf. [Utas 2013]). What was a way of building definite noun phrases in Old Persian, became the standard pattern in Middle Persian and practically the only option from Classical Persian onward. Yet interference from Aramaic and Arabic hardly helps to explain the origin of the *ezāfe* particle, while Elamite and Persian additionally share the SOV and Noun-Genitive word order pattern. Therefore, a gradual language shift from Elamite to Iranian in southwestern Iran must be regarded as the primary reason for the rise of the *ezāfe* construction.

---

[10] This claim is substantially different from the earlier hypothesis of [Gershevitch 1979], according to which Achaemenid Elamite represented a sort of rebus writing, to be always read in Old Persian. See [Tavernier 2008: 75–76] and [Kudrinski, Yakubovich 2016: 55–58] for the critique of Gershevitch's view.

[11] A research work summarizing all the instances of contact-induced developments in scribal milieus still remains a desideratum, but one case that received recently a thorough treatment is that of the Luwian influence on New Hittite [Yakubovich 2010a]. Although one has compelling reasons to reconstruct a progressive shift from Hittite to Luwian in Hattusa and the surrounding area in the 13th century BCE, Hittite remained by tradition the primary written language of the cuneiform chancery. Numerous interference features observed in New Hittite bear witness to the imperfect acquisition of the Hittite written language by Luwian native speakers, just as the interference features of Neo-Elamite and Achaemenid Elamite texts find explanation in the imperfect acquisition of the Elamite written language by Iranian native speakers. Another parallel, namely Buddhist Hybrid Sanskrit, is arguably less precise, since the original written transmission of the relevant texts is not self-evident. For a likely similar scenario concerning Younger Avestan, cf. Section 5.

4. Urartian influence?

A rather different picture emerges when one attempts to trace Urartian influence in the northwest of the Iranian-speaking area, which is invoked in [Windfuhr 2009: 28], alongside the Elamite influence on Old Persian, as a factor in syntactic restructuring (see the beginning of Section 3). Urartian, a language transparently related to Hurrian but lacking other close relatives, is attested in writing in the Kingdom of Van (Urartu) in 9th–7th centuries BCE. Urartian compositions almost exclusively belong to the public domain, royal inscriptions constituting the most widely attested genre. The territory of the Kingdom of Van in the early first millennium BCE roughly corresponds to the area where the Classical Armenian language was in use a millennium later. Scholars tend to agree that already in the period of Urartian literacy part of the population of the Kingdom of Urartu consisted of (Proto-)Armenian native speakers: at least, no hypotheses of Armenian migrations postdating the collapse of the Urartian Kingdom are prominent in the current academic discourse.[12] This suggests that when the Urartian language died out the bulk of Urartian speakers probably shifted to Armenian. Therefore, Armenian morphosyntax is the most obvious place to look for effects of partial restructuring under the influence of a hypothetical Urartian substrate.

Such effects, however, are not immediately obvious. The Classical Armenian language lacks rigid constraints on surface syntax, which enabled, for example, the translation of the Greek Bible into Armenian without syntactic permutations in most cases [Schmitt 2007: 158]. East Armenian, the group of dialects that was historically spoken in Iran and now underlies the official language of the Republic of Armenia, has generalized the left-branching word order in the noun phrase. In this respect, it is similar to the bulk of the Iranian languages, but not to Persian or Kurdish. There could be several reasons why the Urartian speakers failed to leave a long-lasting imprint on the Armenian syntax, but one possible explanation is the modest size of the Urartian language community. This would also explain why Urartian literacy failed to spread to the private sphere and why we have no evidence for either oral or written transmission of the Urartian language in the Achaemenid period. Perhaps the collapse of the kingdom of Van triggered a quick assimilation of the former Urartian elites to the numerically dominant Proto-Armenian population.

What has been said thus far does not bode well for the possibility of intensive Urartian-Iranian bilingualism. Nevertheless, there is one instance where a group of Urartian noun phrases have apparently been calqued into Iranian. This is the chain of royal titles, typical of Ancient Near Eastern "Great Kings". The titles in question ultimately go back to Mesopotamia, and some even find parallels in Egypt, but only the kings of Van and the Achaemenid rulers share their full set, albeit with minor variations (cf. [Schmitt 1977: 386–387, with further references]). Furthermore, the borrowing of Urartian titles in Iran appears to be plausible from the historical viewpoint, given that the Kingdom of Van extended its territory to the south of Lake Urmia in the 8th century BCE and thus became relevant to the Iranian nomads in the area. It is likely, although not strictly provable, that the Urartian set of titles was first appropriated by the Median kings, the erstwhile lieges of Elam/Persia, and then appropriated by the Persian rulers when the balance of power in Iran shifted in their favor under Cyrus the Great. Needless to say, such a type of borrowing does not imply intensive language contact, merely a degree of mutual understanding among the elites. The parallelism between the Urartian and Achaemenid titles is illustrated below.

(19)   ᵐargišti-ni   MAN   *DANNU*   MAN   KUR   šura-we
        Argishti-DEF   king   strong   king   land   all-GEN.PL

---

[12] I concur with the scenario of Proto-Armenian migrations from the west in the wake of the collapse of the Empire of Hattusa, as argued in [Diakonoff 1968]. An alternative hypothesis, implying a much longer presence of Proto-Armenians in the Armenian Highlands, can be found in [Gamkrelidze, Ivanov 1984: 890–894; 898–900]. There is limited evidence for Proto-Armenian function words borrowed into Urartian [Diakonoff 1992; Yakubovich 2010b: 164].

MAN   ^{KUR}biaina-we   MAN   MAN.MEŠ-we
king   Urartian-GEN.PL   king   king-GEN.PL

'Argishti, Strong King, King of all Lands, King of the Urartians, King of Kings…' (A 11-4: 11–12, [Salvini 2008: 542]).

(20) adam   Dārayavauš   xšāyaϑiya   vazarka   xšāyaϑi[ya
    I.NOM   Darius.NOM.SG   king.NOM.SG   great.NOM.SG   king.NOM.SG

    xšāya]ϑiyānām   xšāyaϑiya   Pārsai   xšāyaϑiya   dah[yūnām]
    king.GEN.PL   king.NOM.SG   Persia.LOG.SG   king.NOM.SG   land.GEN.PL

'I am Darius, Great King, King of Kings, King in Persia, King of Lands' (DB. I: 1–2, [Schmitt 2009: 36–37]).

One can see that the morphosyntactic patterns of the titles 'King of Kings' and 'King of (all) Lands' display the same structure Noun-Genitive in both languages. Since the developed case system of Urartian also includes the genitive case, the rendering of the Urartian construction into Old Persian was straightforward, except for the unusual word order. What sets these titles apart from other right-branching noun phrases in Old Persian is the absence of the linker *haya-*/*taya-*. One possible explanation of this discrepancy is the immediate Median origin of the respective formulae [Kent 1953: 95]. The Median language, spoken in central Iran, presumably has not experienced intensive contacts with Elamite. Alternatively, one can hypothesize that the titles under discussion had been borrowed into Persian before the grammaticization of the *kāra haya manā* construction. Whatever their historical origin, they must have retained the flavor of acrolect expressions, similarly to the English titles *heir apparent* or *envoy extraordinary*.[13]

It is under the prism of this comparison that one should consider the origin of the *ezāfe* construction in Kurdish and Zazaki, the Iranian languages that are now spoken on the former territory of the Kingdom of Van. As mentioned above in Section 1, a peculiarity of these languages (or many of their dialects) is that the *ezāfe* linker in right-branching noun phrases can still be inflected for gender, number, and case, as was the situation in Old Persian. We shall illustrate this state of affairs with the case of Zazaki. Example (21) features the oblique masculine and plural linker *-dē*, which contrasts, for example, with direct masculine linker *-ē*, direct feminine linker *-ā*, and oblique feminine linker *-dā* [Paul 2009: 563]. Another peculiarity of the same language is the preservation of the oblique nominal maker *-ī*, which is presumably cognate with the Old Persian gen.sg. ending *-ahyā*. Yet another archaic feature of Zazaki is the *ezāfe* construction with head-noun ellipsis, as in (22).

(21) hētē   nē   'esker-ān-dē   pādīšāh-dē   bīn-īy-ā   rem-en-ā
    to   these   soldier-OBL.PL-LNK.OBL.PL   king-LNK.OBL.M   other-OBL.M-PSTP   run-PRS-3SG.F
    'It runs to the soldiers of the other king' [Paul 2009: 566].

(22) no   lāžek   ē   nāhmān   beg-ī   nī-y-o
    this.M   boy   LNK.NOM.SG.M   Nahman   Beg-OBL   not-be-3SG.M
    'This boy is not (the one) of Nahman Beg' [Paul 2009: 567].

It is probably fair to say that the *ezāfe* construction in Zazaki exhibits more similarities to the *kāra haya manā* in Old Persian than the *ezāfe* construction in New Persian. At the same time, Zazaki, a north-western Iranian language, is no more closely related to Persian than, say,

---

[13] A piece of independent evidence yielding further support to the Urartian origin of the construction under discussion is the Old Persian phrase *vašnā Ahuramadāha* 'by the will of Ahuramazda', which frequently occurs in the Bisitun inscription and also displays right-branching without a linker. It is presumably inspired by the Urartian expression 'by the greatness/might of Haldi', which frequently occurs in the Urartian royal annals but has no obvious counterpart in the Assyrian annalistic tradition [Klein 1988: 393–394, fn. 12, with further references]. In this case, however, the hypothesis of transfer is arguably more complex, because it requires the assumption of oral annalistic tradition in Iran as a source of inspiration for the Bisitun inscriptions.

Parthian or Balochi. One can, therefore, envisage three scenarios that can account for this typological similarity. Either it is purely fortuitous, or it arose in response to similar factors in distinct geographic regions, or the ancestors of Zazaki and Old Persian belonged to the same linguistic area at some point in time.

The first scenario faces the typological problems that have already been addressed above. Zazaki, like Persian, is an SOV language, while the coexistence between SOV and head-dependent word order in the noun phrase is cross-linguistically uncommon. The fact that the same rare correlation is shared by several genetically unrelated languages spoken roughly in the same region (e.g. Akkadian, Elamite, Urartian, and Persian) calls for an explanation in terms of language contact. The second scenario, possibly implied in [Windfuhr 2009: 28], suggests that the ancestors of Kurdish and Zazaki experienced Urartian influence at the northern end of the Zagros tier. Yet Urartian emerges as an unlikely substrate triggering the rise of the *ezāfe* construction, both because it probably had died out early and, more importantly, because no *ezāfe* construction emerged in the sole plausible case of an Urartian phrase calqued by Old Persian. Nor would such a process be likely on formal grounds, because the Urartian possessive constructions, unlike their Elamite counterparts, lack morphemes that could be plausibly calqued as *ezāfe* linkers.

The obvious problem of the third scenario is the geographic distance between the regions where Zazaki is spoken now and where Old Persian was presumably spoken 2500 years ago. Zazaki, the westernmost Iranian language, is at home in central-eastern Turkey, while the Persian heartland is the present-day Iranian province of Fars to the northeast of the Persian Gulf. Yet the gap of almost 2000 km between these two regions can be bridged: the Kurdish dialects, which occupy most of the interim space, share the *ezāfe* construction. Furthermore, it is likely on independent grounds that the Kurdish and Zazaki expansion to the northeast postdates the Achaemenid period. The sheer similarity of the Kurdish dialects implies the date of no earlier than the first millennium CE as the starting point of their filiation. The phonological innovations of the northern group of dialects (Kurmanji) suggest the distinct impact of the Armenian substrate [Tsabolov 1997: 7–8], which tips scales in favour of localizing the Kurdish homeland in the south of the present-day Kurdish spread zone.[14] The large number of phonological isoglosses shared by Persian and Kurdish, which set Kurdish apart from the other Northwest Iranian languages [Windfuhr 2009: 19–20], supports the same conclusion.[15]

As far as Zazaki is concerned, the analysis of its dialectal profile suggests that the speakers of this language were at home in northern Iran and their migrations to northern Mesopotamia and

---

[14] It seems worth stressing once again that the question about the linguistic homeland of the Kurds should be treated separately from the politically charged issues of the origin of the ethnonym *Kurd* or the cradle of Kurdish national identity. In particular, the inferences made here remain perfectly compatible with the observation that the expression "land of the Kurds" is applied to various areas of Upper Mesopotamia in its earliest attestations in Syriac sources. One should keep in mind that not all the varieties of Kurdish are mutually intelligible, and therefore, sociolinguistic matters aside, we are rather dealing with several languages belonging to the Kurdish linguistic area.

[15] Cf. the following formulation: "The fact that some of the innovations which have occurred in the various stages of Persian have been taken over by Balochi (and some more by Kurdish) indicates that Balochi and Kurdish have been influenced by Persian since Old Iranian times" [Korn 2003: 59]. At the same time, one of the anonymous reviewers informs me that the old areal isoglosses with Persian are more strongly felt in Northern Kurdish (Kurmanji) than in Central Kurdish dialects. Compare, for example, Kurmanji *čār* 'four' vs. Central Kurdish *čār*, *čuwār* 'id.' [Tsabolov 2001–2010, I: 229] or Kurmanji *dil* 'heart' vs. Central Kurdish *dil*, *dir*, *zir* 'id.' [Tsabolov 2001–2010, I: 307], where the Kurmanji forms are close to Classical Persian *čahār* 'four' and *dil* 'heart' respectively. Since the Kurmanji speakers are rarely bilingual in Persian, whereas the speakers of Central Kurdish dialects frequently use Persian is their second language, this discrepancy does not find explanation in terms of recent contacts with Persian. One can rather envisage the scenario of the Central Kurdish dialects amalgamating residual forms of Northwest Iranian vernaculars as a result of recent language shifts, but this question, of course, requires further study.

then Anatolia do not predate the Parthian period [Paul 1998]. There is hardly any doubt that the coexistence of the OV and N-Gen basic word order patterns in Zazaki shows the impact of language contact. On the one hand, the languages of northern Iran, the homeland of Zazaki, typically exhibit consistent left-branching (cf. Section 1), on the other hand, vestiges of the original left-branching nominal syntax are the Zazaki postpositions, such as *bin* 'under', *ser* 'on(to)' etc. [Paul 2009: 553]. Since the Kurdish language is recognized as the most important source of influence on Zazaki [Paul 2009: 546], it is logical to attribute the new right-branching dominant pattern of Zazaki nominal syntax to the Kurdish (Kurmanji) superstrate. The restructuring in certain North-Eastern Aramaic dialects, which yielded similar discrepancy between word order patterns under Kurdish impact, supports the same conclusion, even though in this case the syntactic change primarily concerned object-verb placement [Haig 2015: 409–413].

Therefore, it seems reasonable to treat the spread of the *ezāfe* construction to the northwest, across the Zagros Mountains and the present-day borders of Iran, as a relatively recent phenomenon, which accompanied the latest wave of Iranian westward migrations. Given that the Kurdish language is not attested in written sources up to the 16th century CE, while the Zazaki language came on the radar in the mid-19th century, the sociolinguistic details of this process will remain a matter of speculation. For example, one cannot be sure whether the speakers of Proto-Kurdish borrowed the construction under discussion from the Old Persian superstrate/adstrate or directly from the Elamite substrate. The last possibility, however, is by no means ruled out: even today the Iranian province of Ilam, situated immediately to the north of the former Elamite capital Susa, has a predominantly Kurdish population. Likewise, one cannot state with certainty, whether the Zaza migrations to the west through the Kurdish-speaking territory were accompanied by the shift of Kurdish speakers to Zazaki (Restructuring I) or stable bilingualism with Kurdish functioning as the acrolect (Restructuring II). But whichever scenario one favors, the right-branching *ezāfe* construction in Zazaki must ultimately be due to Kurdish influence.

Summing up, despite the formal similarities between the syntax of Urartian and the Iranian languages spoken in the same area nowadays, the Urartian influence on Kurdish or Zazaki cannot be proven. Accordingly, there is no reason to treat Urartian as an Iranian substrate on a par with Elamite. Occam's razor supports an alternative scenario of the recent spread of the *ezāfe* construction to the Armenian Highlands. A practical lesson to be learned from this discussion is the necessity to draw on both typological and historical expertise when tackling language contact in ancient societies.

5. Special cases

The goal of this section is to address the data that represent potential challenges to the proposed scenario of language contact. There are several types of noun phrases that partially resemble the *ezāfe* construction but occur in several Old and Middle Iranian languages that never experienced direct contact with Elamite. These are Bactrian appositional constructions, possessive noun phrases in Khwarezmian, and right-branching construction with the linker *ya-* in Younger Avestan. While the exhaustive description of these syntactic units cannot be attempted here, I intend to demonstrate that the first two units are formally different from the *ezāfe* construction in Persian, while the situation in Younger Avestan probably reflects secondary restructuring under Persian influence.

The Bactrian language of what is now northern Afghanistan lost its nominal case endings by the early first millennium CE. Therefore, the inherited preposition of syntactic dependents constitutes the main device of organizing Bactrian noun phrases, e.g. μασκο ναβιχτιγο οιγαλφανο 'above-written witnesses', μο λαϸνοβωστογο 'this gift deed', σασανο πορο 'Sasan's son'. A different construction is, however, deployed for expressing syntactic apposition, usually in combinations of names and titles. In this case, the most salient part of the phrase comes first, while the following elements are frequently, although not always, linked by the particle ι [Gholami

2011: 17]. Such examples, a representative selection of which is provided in Table 2, bear a degree of resemblance to the *ezāfe* construction, e.g. Middle Persian *Ardašīr ī Pābagān* 'Ardashir, son of Pabag'. In fact, Nicholas Sims-Williams, the decipherer of the Bactrian administrative texts, uses the term *ezāfe* with reference to this particle and tentatively connects it with the Iranian relative pronoun **ya-* [Sims-Williams 2007: 214b; 2009: 261]. The construction with the linker ι is, however, never used for the expression of possession in Bactrian [Gholami 2011: 17].

Table 2

Bactrian ι used as linker

(all examples from [Sims-Williams 2000; 2007])

Source	Example	Translation
A9–10	νινδοκο ι ωχþοβαδογανο	'Ninduk Okhshbadugan'
G2–3	μοζ[δο ι χαρα]γανο ι κανδογοληρο	'Mu[zd Khara]gan, the granary-keeper
G3–4	ω[ρομ]ο̅ζ̅δ̅ο̅ ι βο[ρνικα]ιο	'O[rm]uzd Bu[rnika]n'
Ii5	ραμογολο ι ζινδ[οκο πορο]	'Ram-gul [son of] Zind[uk]'
Ii6	[ωρο̅]μ̅ο̅ζ̅δο πιδοοιησιμο ι σασανο πορο	'(I) [Or]muzd declare, the son of Sasan'
L2	βαγο ι οαχþο	'god Wakhsh'
al18	οιριþτομιþο ι χοη[οι]	'Wirishtmish the lord'
bb3	[μ]τ̅ρ̅ο̅μ̅α̅[ρη]γο φρομαλαρο ι μ[α]ρηγο	'Mihr-ma[re]g the steward, (your) servant'

There is, furthermore, another group of cases, collected in Table 3 below, where Bactrian ι functions as an optional definite article (a proclitic variant of the pronoun ειο 'this'). In such cases, [Sims-Williams 2007: 214b] admits the connection of ι with the Iranian proximal demonstrative pronoun **ayam*, which is also accepted in [Gholami 2011: 11]. Given the formal possibility of such an etymology, it seems equally plausible that **ayam* also represents an archetype of the linker ι in Table 2. There are many languages, including English, where the definite article can function as linker in the appositional construction, e.g. *Richard the Lionheart*, *Jude the Obscure*. If one accepts a similar semantic extension for Bactrian, then the parallelism between the expression of apposition in Bactrian and Persian turns out to be a mere coincidence. Alternatively, given that Bactria belonged to the Sasanian Empire in the 3rd century CE, one cannot exclude contact-driven convergence with the Middle Persian onomastic formulae.

Table 3

Bactrian ι used as article

(all examples from [Sims-Williams 2000])

Source	Example	Translation
A3	ι ωνδο λιβο	'the present document'
B4	σιδασο ι ναμαγο γιβισιδο	'what is written in the list'
G5	ασο ι χοβο οαρ[ζο]	'from the personal prod[uct]'
I5	αβο ι ηβ[οδαλαγγο τωγο]	'toward the He[phthalite tax]'
J14–15	ι μασκο ναβιχτιγο αγγαργο	'the property described therein'
L6	ι βονο καδγο	'the primary residence'
L11	αβο ι παρο	'toward the debt'
L18–19	ι πανδαγο ασιδο οαρο ναυαγγινδο	'the path that is adjacent there(to)'

A complication of the proposed analysis is the presence of the rare feminine form ια 'this'. On face value, it supports the contamination with the relative pronoun, since the feminine counterpart of Iranian *ayam 'this' was *iyam, which could not yield ια in Bactrian through a regular sound change. Yet one cannot exclude the early analogical replacement of *iyam with *ayā or *iyā (cf. the discussion of Khwarezmian below). Even if *yā, nom.sg.f of *ya-, represented an additional source of analogy, there is no gain in restricting its scope to the secondary appositional construction. To be sure, the Bactrian language partook in the functional extension of the interrogative pronoun *ka-/či- at the expense of the old relatives, in accordance with the common pattern of Middle Iranian. But the relative pronoun *ya- clearly existed in the ancestor of the Bactrian language at some point, and could accordingly influence the paradigm of the proximal demonstrative.

This having been said, there remains a formal possibility that the Bactrian appositional construction had generalized the relative pronoun *ya- before it merged with the demonstrative pronoun *ayam. Even under such an assumption, there is no reason to treat the Bactrian morphosyntactic change as a close parallel to the rise of the *ezāfe* construction in Persian. Crucially, the Bactrian development did not involve a change in the direction of branching. Already in Old Iranian, the epithets and other appositions tended to follow the elements they modify, e.g. Y. 29.8 *Zaraϑuštrō Spitāmo* 'Zarathushtra Spitama', Y. 32.3 *yūš daēuuā vīspåŋhō* 'you, demons of all kinds' (nom.) [West 2011: 10–11]. A linker could occasionally be inserted to reinforce this construction, e.g. Y. 32.5 *vå̊ ... yə̄ŋg daēuuə̄ŋg* 'you, who (are) demons' (acc.). In essence, the head / dependent assignment in appositional constructions is the matter of pragmatics rather than grammar, and therefore the word order in such phrases can hardly be seen as typologically informative.

Now we can move to the situation in Khwarezmian, an Iranian language that was spoken immediately to the south of the Aral Sea up to the Early Islamic period. The Khwarezmian definite article, *ī* in the masculine and *yā* in the feminine, was not declined for case. According to [Durkin-Meistererent 2009: 343], it "patently originates in the Old Iranian relative pronoun", but such a functional development is typologically quite uncommon. As one option, one can treat the Khwarezmian article as a result of contamination between the demonstrative pronoun *iyam and the relative pronoun *ya- (nom.sg.f *yā). This is essentially the same solution that was discussed above in connection with Bactrian. Alternatively, one can assume that the demonstrative pronoun *ayam built the secondary form *ayā (or *iyā) through leveling to *tā 'that (nom. sg.f)' and perhaps to the feminine adjectives in *-ā. Both of these possibilities have already been addressed in [Edel'man 2008: 30–31].

What makes the situation in Khwarezmian distinct from Bactrian is the innovative right-branching word order, which extends itself to all sorts of noun phrases, not just to the appositions. In the instance of the nominal and adjectival constituents, both the head and its right-branching dependent are accompanied by the article. The postposed adjective always takes the article of the head noun, while the article of the postposed genitive agrees with it in gender, as in (23). The possessive forms in /-ān/ display the syntactic behavior of adjectives rather than genitives. Thus, in the instance of (24–25), the gender of the noun *w'c* 'matter' is irrelevant for the formation of the possessive construction, since the article that precedes it merely duplicates that of the head noun. All the examples below are taken from [Durkin-Meistererent 2009: 358].[16]

(23) 'y   b'r   y'   'wdry
      ART.SG.M load.NOM.SG   ART.SG.F belly.GEN.SG(F)

      'the load of the belly (= foetus)'

---

[16] It is worth noting in passing that the morpheme /-ān/, present in *w'c-'n* and traditionally analyzed as a possessive case marker, may well go back to the adjectival suffix *-āna [Edel'man 2008: 28]. If so, it is the same morpheme that yielded the marker of Middle Persian and Bactrian patronymics. For the occurrence of Bactrian patronymics as second members of the appositional construction, cf. Table 2 above.

(24) ’y   ḥqyqt   ’y   w’c’n
      ART.SG.M   truth.NOM.SG   ART.SG.M   matter.POSS
      ‘the truth of the matter’

(25) y’   pc   y’   w’c’n
      ART.SG.F   end.NOM.SG   ART.SG.F   matter.POSS
      ‘the end of the matter’

Examples (24)–(25) and the like may appear akin to the Persian *ezāfe* construction, and were indeed treated as suggestive evidence for the contamination between the demonstrative and relative pronouns yielding the Khwarezmian article [Edel'man 2008: 31, with further references]. Such a reconstruction is certainly possible, but this is not the only possibility. Another crucial innovation of the Khwarezmian language is the rise of the verb-initial word order [Ibid.: 54–57], which was accompanied by the head marking of the nominal arguments with the help of clitic pronouns, also known as clitic doubling (for the relevant examples see [Durkin-Meisterernst 2009: 360–361]). This phenomenon, quite unusual as it is for the Iranian languages, provides nonetheless a perfect correlation for the noun phrase word order in (23)–(25): the VSO/VOS patterns and the head-dependent order in the noun phrase are both typical manifestations of right-branching. Whether or not the ongoing radical inversion of word order in Khwarezmian was substrate-induced remains an open question, but even if it was, this putative substrate has nothing to do with Elamite, which is characterized by verb-final syntax.

Neither is there any necessity to treat the repetition of articles, deployed as a linking device in (24)–(25), as a close genetic or areal match of the *kāra haya manā* construction in Old Persian. If anything, both types of noun phrases reflect a cross-linguistic tendency toward the rise of morphosyntactic complexity in the transition from left to right-branching (cf. e.g. English *the enemy lines* vs. *the lines of the enemies*). A closer parallel to (24)–(25) would be the phrase οἱ ἄνδρες οἱ ἀγαθοί ‘good men’, which was available in Attic Greek as a secondary and formally more complex alternative to οἱ ἀγαθοὶ ἄνδρες ‘id.’ In a sense, both the Greek and Khwarezmian right-branching constructions may be regarded as instances of mismatch between the attributive/possessive semantics and appositional syntax. A test that confirms such an interpretation for Khwarezmian is the possibility to repeat prepositions in the situation of right-branching: *f-y zβ’k f-y trk’nk* ‘in the tongue of the Turks’ [Durkin-Meisterernst 2009: 358]. Thus, although contamination with the relatives remains theoretically possible, the derivation of articles from demonstrative pronouns appears to be fully sufficient for a typologically informed account of the Khwarezmian noun phrases under discussion.

The case of Younger Avestan is quite different. Here one encounters numerous noun phrases where syntactic dependents are assuredly linked to their heads by means of relative pronouns or particles, which is the exact archetype of the Persian *ezāfe* construction. A typical example is a passage from the ritualistic part of the Avesta known as the *Vidēvdād*: *daēnąm yąm āhūirīm zaraϑuštrīm* ‘the religion of Ahura and Zarathushtra (acc.)’ (Vd. 2.1). Yet the paradigm of the linker is irregular in an interesting way. To begin with a minimal pair, in Vd. 2.1 we encounter a contrast between the nominative and dative appositional constructions *tūm yō ahurō mazdå* ‘thou, Ahura Mazda (nom.)’ and *mana yaṯ zaraϑuštrāi* ‘to me, Zarathushtra (dat.)’. In the first case, the regular nom.sg.m form of the relative pronoun is in use, in the second one the relative pronoun is replaced with the particle *yaṯ*, historically a form of nom.-acc.sg.n.

In general, nominative, accusative, and instrumental forms of the relative pronoun in the linking function are mostly preserved, while the other case forms disappear in favor of *yaṯ* in Younger Avestan [Haider, Zwanziger 1984: 149 with further references]. This mismatch can be additionally illustrated by the ablative phrase Yt. 17.19 *haca zəmaṯ yaṯ paϑanaiiå* ‘from the broad Earth’, the locative phrase Vd. 5.39 *ahmi aŋhuuō yaṯ astuuaiṇti* ‘in this material existence’ or the genitive phrase Vd. 3.15. *aētahe narš … yaṯ iristō.kaṣ̌ahe* ‘of that man (who is) an undertaker’. It is compatible with the hypothesis that the *ezāfe* construction in Younger

Avestan represents an attempt to copy a similar construction from another language, featuring a more restricted paradigm of grammatical cases. A possible reason for the preservation of the declined linkers in the accusative and instrumental phrases is the presence of the Old Avestan prototypes for the relevant constructions, cf. (14) and (15) above. If this is a viable explanation, the contact-induced construction was selected in the course of second-language acquisition of Avestan, precisely for those instances where the Old Avestan corpus failed to provide a ready-made pattern.[17]

The obvious candidate for the trigger of this process is Persian, the language (or rather succession of languages) that was particularly likely to affect the transmission of the Avesta from the sociolinguistic perspective. Already in the Achaemenid Empire, the magi-priests, described by Herodotus (1.140) as devout Zoroastrian practitioners, were integrated into the administration of the state cult [Dandamaev 2012]. In the Sasanian period, Zoroastrianism became the state religion of Iran, and the first written fixation of the Avesta was presumably accomplished under the patronage of the Sasanian rulers.

One must, however, keep in mind that the Old Persian language of royal inscriptions restricts the *kāra haya manā* construction to the nominative, genitive and ablative-instrumental cases (cf. Section 2). This begs the question about the variety of Persian that could impact the spread of the the *ezāfe* construction to other cases in Younger Avestan. In my opinion, this was the lingua franca on the Achaemenid provincial elites, which can be called for convenience Peripheral Old Persian, and which anticipated the evolution of Late Old Persian, as described in [Schmitt 1999: 59–118]. The drastic simplification of Persian morphology over some two centuries of the existence of the Achaemenid Empire can be best accounted for by assuming the massive shift to Persian on the part of its subjects, inevitably accompanied by imperfect learning. The nominal declension of Peripheral and Late Old Persian was presumably reduced to the opposition between the direct and oblique cases, with no gender distinctions. In contrast, the *ezāfe* linker, which lacked the inherited genitive-dative form, may have been reduced to a single particle, which could then easily spread across the board.[18] The learners of Younger Avestan would accordingly extend their *ezāfe* construction to all types of noun phrases but use the inherited form *yaṯ* as the equivalent of the Persian *ezāfe* particle in secondary constructions.

Another approach to investigating the *ezāfe* construction in Younger Avestan is the study of metric irregularities. Certain liturgical parts of the Avesta are written in eight-syllable lines, which resemble the basic units of the Sanskrit *ślokas* and therefore may well reflect Indo-Iranian heritage. While the Younger Avestan meter notoriously features many exceptions from this basic rule, some of the surface irregularities may be eliminated by way of restoring the original text, unaffected by the corruptions of the subsequent oral or written transmission. Below, such an experiment is applied to the passage Y 10.1 from a hymn to Haoma (cf. [Kellens 2003] and [Pirart 2004] for its philological treatment and translation). One can see in the last two lines that part of the price for obtaining a regular metric structure is assuming a secondary origin of the *ezāfe*-like particles.

---

[17] An alternative account, offered in [Reichelt 1909: 370], explains this distribution in syllabic terms: the monosyllabic forms of the relative pronoun were preserved in the linker function, while the disyllabic forms were replaced with the particle. Such an analysis is certainly possible but not mutually exclusive with the one proposed here: the constraint on monosyllabic linkers may have been the fact of Old Avestan, where linkers were generally fairly rare. After they gained in frequency, the *ezāfe* construction was generalized to all the cases with the help of the new particle, but this already happened within the context of liturgical language acquisition.

[18] A testimony to the innovative *ezāfe* particle *t-y* in Late Old Persian is A2Sc. 5-6 *i-m-a-m* [*u-s-t-*]*c-n-a-m t-y a-θ-g-i-n-a-m* 'this stone staircase' [Schmitt 2009: 194], where *t-y* is used instead of the expected *t-y-a-m*. This example was most recently compared with the Younger Avestan *ezāfe* construction in [Meyer 2017: 216].

Table 4

Metric restoration of Y. 10.1

Normalized Text	Restored Text (8-syllable meter)
viša apąm iδa patəṇtu	wiš' apām ida patantu
vī daēuuåṇhō vī daēuuaiiō	wi daiwāhah wi daiwayah
vaŋhuš sraošō mitaiiatu	wahuš sraušah mitāyatu
ašiš vaŋᵛhi iδa miϑnatu	artiš wahw' ida miϑnātu
ašiš vaŋᵛhi rāmiiaṯ iδa	artiš wahwī rāmyāt ida
upa imaṯ nmānəm yaṯ āhūiri	up' imat dmānam <yat> ahuri
yaṯ haomahe ašauuazaŋhō	<yat> haumahya artawazahah

What justifies the observations and experiments of the type outlined above is the presumable transmission of Avestan in a multilingual or non-native environment for a long period of time. The Avestan texts are limited to Zoroastrian religious literature. The earliest of them, written in Old Avestan, reflect the native language of the founders of the new religion, arguably including the prophet Zarathushtra. At that time, the language must have been fairly close to common Iranian, since it displays no genetic innovations characterizing West Iranian and no undisputable areal innovations that are associated with East Iranian.[19] The other Avestan texts either arose within the fold of Zoroastrian tradition or represent the adaptation of pre-existing Iranian texts to this tradition. While the Avestan language may have lingered on as a vernacular for some period of time, it has no modern direct descendants. The conservative account of [de Vaan 2003: 14] suggests the approximate date of 300 BCE for the extinction of Avestan as a living language, while the most commonly held date for the written fixation of the Avesta is some six to eight hundred years later [Hoffmann, Forssman 2004: 36].

The Persian influence on Avestan triggering the spread of the *ezāfe* construction may have taken various forms. Since the majority of the scholars assume that the compilation of the *Vidēvdād* dates back to the post-Achaemenid period [Malandra 2006], the whole composition may have come into being in the bilingual circles of the *magi*-priests. If so, the right-branching noun phrases with the linker *ya-* in the *Vidēvdād* and several other late compositions may be summarily treated as calques of the *kāra haya manā* construction in Old Persian or rather its immediate descendant.[20] In sociolinguistic terms, such restructuring in a non-natively transmitted language resembles the development of Achaemenid Elamite, except for the original oral transmission of the *Vidēvdād*. In the instance of several Avestan hymns to Zoroastrian deities, the pre-Achaemenid date of the respective compositions appears more likely. Nevertheless, linkers could still be secondarily inserted in the process of their subsequent transmission, even at the expense of corrupting their metric structure, as in the example in Table 4. The absence of similar

---

[19] For East Iranian as a linguistic area rather than a genetic branch of the Iranian languages, see [Sims-Williams 1989: 165]. In terms of linguistic geography, the contrast between the status of East and West Iranian correlates with the assumption that the speakers of the East Iranian languages stayed in and around the area of the postulated Central Asian homeland of the Iranians, whereas the West Iranian speakers migrated farther westward to the Iranian Plateau.

[20] The possible impact of non-native transmission on the morphosyntax of Younger Avestan rarely attracts scholarly attention, perhaps because the majority of scholars working on the Avestan grammar were trained as Indo-Europeanists rather than sociolinguists. While the comprehensive description of the relevant Younger Avestan innovations clearly falls beyond the scope of this paper, an additional morphosyntactic phenomenon that is unlikely to have come into being within a community of native speakers is the secondary use of instrumental plural forms for Younger Avestan subjects and objects, most recently studied in [de Vaan 2018]. The smoking gun in this case is the possible lack of morphological agreement in a noun phrase, e.g. Y 15.1 *vaŋhūš srīrāiš nāmąn* 'good beautiful names', where only the second word has the instrumental ending. Presumably such hypercorrections occurred as a reaction to a vernacular where dedicated instrumental plural forms were no longer in use.

insertions in Old Avestan texts has to do with their particularly sacred character within the tradition: their final shape was fixed earlier and transmitted rather more carefully.[21]

To be sure, even if one strips the Younger Avestan corpus of all the Achaemenid and post-Achaemenid additions and secondary insertions, one will wind up with the residue of right-branching constructions with pronominal linkers in Younger Avestan, for which the explanation in terms of language contact does not impose itself. The quantification of this phenomenon must remain a task for the future, but in notional terms it can be compared with the situation in Parthian and particularly Old Avestan, where such constructions existed but played a subordinate role. The restricted use of verbless relative clauses modified by case attraction, the ancestors of the later *ezāfe* construction, can be arguably reconstructed for common Iranian (cf. Section 2). In contrast, the generalization of such a construction as the primary or one of primary means of organizing the noun phrase is restricted to a fairly specific set of Iranian languages. These are languages where one can assume the presence of the Elamite substrate, such as Old Persian and perhaps the ancestor of Kurdish, modern descendants of such languages, such as Farsi or Tajik, and languages affected by secondary contact-induced restructuring, such as Zazaki or Younger Avestan. The rest of the Iranian languages lack the construction under discussion, while its assumed manifestations in Bactrian and Khwarezmian receive alternative explanations.

GRAMMATICAL ABBREVIATIONS

ABL — ablative
ACC — accusative
ANIM — animate
ART — article
DEF — definite
F — feminine
GEN — genitive
IMPF — imperfect

INAN — inanimate
INSTR — instrumental
LNK — linker
LOC — locative
M — masculine
N — neuter
NOM — nominative
OBL — oblique

PERF — perfect
POSS — possessive
PL — plural
PRS — present
PRT — preterit
PSTP — postposition
REL — relative
SG — singular

PHILOLOGICAL ABBREVIATIONS

A2Sc. — Artaxrexes II, Staircase
DB. — Darius I, Bisitun
DNa. — Darius I, Naqsh-e-Rostam (a)
SghS. — Saṅghāṭasūtra

Vd. — Vidēvdād
XPf. — Xerxes, Persepolis (f)
Y. — Yasna
Yt. — Yašt

REFERENCES

Adiego 2000 — Adiego Lajara J.-I. Sobre el relative del antiguo persa. *Indoarisch, Iranisch und die Indogermanistik. Arbeitstagung der Indogermanischen Gesellschaft vom 2. Bis 5. Oktober 1997 in Erlangen*. Forssman B., Plath R. (eds.). Wiesbaden: Reichert, 2000, 1–13.

Bae 2001 — Bae C.-H. *Comparative studies of King Darius's Bisitun Inscription*. Ph.D. diss. Harvard Univ., 2001.

Boyce 1964 — Boyce M. The use of relative particles in Western Middle Iranian. *Indo-Iranica: Mélanges présentés à Georg Morgenstierne à l'occasion de son soixante-dixième anniversaire*. Redard G. (éd.). Wiesbaden: Harrassowitz, 1964, 28–47.

---

[21] The secondary insertions corrupting the metric structure of Old Avestan texts are also well known [Gippert 2000: 166–167]. In this case, however, one is dealing not with direct interference effects but rather with exegetic annotation.

Boyce 1975 — Boyce M. *A reader in Manichaean Middle Persian and Parthian*. Leiden: Brill, 1975.

Brunner 1977 — Brunner C. J. *A syntax of Western Middle Iranian*. Delmar (NY): Caravan Books, 1977.

Brust 2018 — Brust M. *Historische Laut- und Formenlehre des altpersischen, mit einem etymologischen Glossar*. Innsbruck: Institut für Sprachen und Literaturen der Universität Innsbruck, 2018.

Canevascini 1993 — Canevascini G. *The Khotanese Saṅghāṭasūtra: A critical edition*. Wiesbaden: Reichert, 1993.

Cathcart 2015 — Cathcart Ch. A. *Iranian dialectology and dialectometry*. Ph.D. diss. Univ. of California, 2015.

Dandamaev 2012 — Dandamaev M. A. Magi. *Encyclopedia Iranica*, online version. New York: Columbia Univ. URL: http://www.iranicaonline.org/articles/magi.

Diakonoff 1968 — Дьяконов И. М. *Предыстория армянского народа. История Армянского нагорья с 1500 по 500 г. до н. э. Хетты, лувийцы, протоармяне*. Ереван: АН Армянской ССР, 1968. [Diakonoff I. M. *Predystoriya armyanskogo naroda. Istoriya Armyanskogo nagor'ya s 1500 po 500 g. do n. e. Khetty, luviitsy, protoarmyane* [Prehistory of the Armenian people. History of the Armenian plateau in 1500–500 BCE. Hittites, Luwians, Proto-Armenians]. Yerevan: Academy of Sciences of the Armenian SSR, 1968.]

Diakonoff 1992 — Diakonoff I. M. First evidence for the Proto-Armenian language in Eastern Anatolia. *Annual of Armenian Linguistics*, 1992, 13: 51–54.

Dryer 1992 — Dryer M. S. The Greenbergian word order correlations. *Language*, 1992, 68: 81–138.

Dryer, Haspelmath 2013 — Dryer M. S., Haspelmath M. (eds.). *World atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology, 2013. URL: https://wals.info/ (accessed 13.08.2019).

Durkin-Meisterernst 2009 — Durkin-Meisterernst D. Khwarezmian. *The Iranian languages*. Windfuhr G. (ed.). London: Routledge, 2009, 336–376.

Durkin-Meisterernst 2014 — Durkin-Meisterernst D. *Grammatik des Westmitteliranischen (Parthisch und Mittelpersisch)*. Vienna: Österreichische Akademie der Wissenschaften, 2014.

Edel'man 2008 — Эдельман Д. И. Хорезмийский язык. *Основы иранского языкознания. Среднеиранские и новоиранские языки*. Ефимов В. А. (отв. ред.). М.: Восточная лит-ра, 2008, 6–60. [Edel'man D. I. Khwarezmian. *Osnovy iranskogo yazykoznaniya. Sredneiranskie i novoiranskie yazyki*. Efimov V. A. (ed.). Moscow: Vostochnaya Literatura, 2008, 6–60.]

Efimov 1975 — Ефимов В. А. Категория рода. *Опыт историко-типологического исследования иранских языков*. Расторгуева В. С. (отв. ред.). Т. 2. М.: Наука, 1975, 7–104. [Efimov V. A. The gender category. *Opyt istoriko-tipologicheskogo osveshcheniya iranskikh yazykov*. Rastorgueva V. S. (ed.). Vol. 2. Moscow: Nauka, 1975, 7–104.]

Gamkrelidze, Ivanov 1984 — Гамкрелидзе Т. В., Иванов Вяч. Вс. *Индоевропейский язык и индоевропейцы. Реконструкция и историко-типологический анализ праязыка и протокультуры*. Тбилиси: Тбилисский университет, 1984. [Gamkrelidze T. V., Ivanov Vyach. Vs. *Indoevropeiskii yazyk i indoevropeitsy. Rekonstruktsiya i istoriko-tipologicheskii analiz prayazyka i protokul'tury* [The Indo-European language and the Indo-Europeans. Reconstruction and historical-typological analysis of the proto-language and proto-culture]. Tbilisi: Univ. of Tbilisi, 1984.]

Gershevitch 1979 — Gershevitch I. The alloglottography of Old Persian. *Transactions of Philological Society*, 1979: 114–190.

Gholami 2011 — Gholami S. Definite articles in Bactrian. *Topics in Iranian linguistics*. Korn A. et al. (eds.). Wiesbaden: Reichert, 2011, 10–22.

Gippert 2000 — Gippert J. The Avestan language and its problems. *Indo-Iranian languages and peoples*. Sims-Williams N. (ed.). Oxford: British Academy; Oxford Univ. Press, 2000, 165–187.

Goedegebuure 2008 — Goedegebuure P. Central Anatolian languages and language communities in the Colony Period: a Luwian-Hattian symbiosis and independent Hittites. *Anatolia and the Jazira during the Old Assyrian period*. Dercksen J. G. (ed.). Istanbul: Nederlands Historisch-Archaeologisch Instituut in het Nabije Oosten, 2008, 137–180.

Greenberg 1963 — Greenberg J. Some universals of grammar with particular reference to the order of meaningful elements. *Universals of human language*. Greenberg J. (ed.). Cambridge (MA): MIT Press, 1963, 40–70.

Grillot-Susini 2008 — Grillot-Susini F. *L'élamite : élements de grammaire*. Paris: Geuthner, 2008.

Haider, Zwanziger 1984 — Haider H., Zwanziger R. Relatively attributive: The 'ezāfe' construction from Old Iranian to Modern Persian. *Historical Syntax*. Fisiak J. (ed.). Berlin: Mouton, 1984, 137–172.

Haig 2015 — Haig G. Verb-goal (VG) word order in Kurdish and Neo-Aramaic: Typological and areal considerations. *Neo-Aramaic and its linguistic context.* Khan G., Napiorkowska L. (eds.). Piscataway: Gorgias, 2015, 407–425.

Henkelman 2011 — Henkelman W. F. M. Cyrus the Elamite and Darius the Persian: A case of mistaken identity. *Herodot und das Persische Weltreich — Herodotus and the Persian Empire: Akten des 3. Internationalen Kolloquiums zum Thema „Vorderasien im Spannungsfeld klassischer und altorientalischer Überlieferungen", Innsbruck, 2008*. Rollinger R.et al. (eds.).Wiesbaden: Harrassowitz, 2011, 577–634.

Hoffmann, Forssman 2004 — Hoffmann K., Forssman B. *Avestische Laut- und Flexionslehre*. 2nd edn. Innsbruck: Universität Innsbruck, 2004.

Isaev 1987 — Исаев М. И. Осетинский язык. *Основы иранского языкознания. Новоиранские языки: восточная группа*. Расторгуева В. С. (отв. ред.). М.: Наука, 1987, 537–643. [Isaev M. I. Ossetian. *Osnovy iranskogo yazykoznaniya. Novoiranskie yazyki: vostochnaya gruppa*. Rastorgueva V. S. (ed.). Moscow: Nauka, 1987, 537–643.]

Ivanov, Dodykhudoeva 2017 — Иванов В. Б., Додыхудоева Л. Р. Синтаксические отношения имен в северо-западных иранских языках (на мазандеранском и гилякском материале). *Вопросы языкознания*, 2017, 2, 77–95. [Ivanov V. B., Dodykhudoeva L. R. Syntactic relations of nominals in the North Western Iranian languages (with special focus on Mazandarani and Gilaki). *Voprosy Jazykoznanija*, 2017, 2, 77–95.]

Jahani, Korn 2009 — Jahani K., Korn A. Balochi. *The Iranian languages*. Windfuhr G. (ed.). London: Routledge, 2009, 634–692.

Kellens 2003 — Kellens J. *Le Hōm Stōm et la zone des déclarations (Y 7.24–Y 15.4, avec les intercalations de Vr3 à 6)*. Paris: De Boccard, 2003.

Kent 1953 — Kent R. G. *Old Persian: Grammar, texts, lexicon*. 2nd edn. Hew Haven: American Oriental Society, 1953.

Klein 1988 — Klein J. S. Coordinate conjunction in Old Persian. *Journal of the American Oriental Society*, 1988, 108/3: 387–417.

Korn 2003 — Korn A. Balochi and the concept of North-Western Iranian. *The Baloch and their neighbours*: *Ethnic and linguistic contact in Balochistan in historical and modern times*. Jahani K., Korn A. (eds.). Wiesbaden: Reichert, 2003, 49–61.

Kudrinski, Yakubovich 2016 — Kudrinski M., Yakubovich I. Sumerograms and Akkadograms in Hittite: Ideograms, logograms, allograms, or heterograms? *Altorientalische Forschungen*, 2016, 43/1–2: 53–66.

Malandra 2006 — Malandra W. M. VENDĪDĀD i: Survey of the history and contents of the text. *Encyclopedia Iranica*, online version. New York: Columbia Univ. URL: http://www.iranicaonline.org/articles/vendidad (accessed 13.08.2019).

Meyer 2017 — Meyer R. *Iranian-Armenian language contact in and before the 5th century CE*. Ph.D. diss. Univ. of Oxford, 2017.

Paul 1998 — Paul L. The position of Zazaki among the Iranian languages. *Proc. of the 3rd European Conf. on Iranian Studies (Cambridge, 1995)*. Part I: *Old and Middle Iranian Studies*. Sims-Williams N. (ed.). Wiesbaden: Reichert, 1998, 163–177.

Paul 2009 — Paul L. Zazaki. *The Iranian languages*. Windfuhr G. (ed.). London: Routledge, 2009, 545–586.

Paul 2013 — Paul L. *A Grammar of Early Judaeo-Persian*. Wiesbaden: Reichert, 2013.

Pirart 2004 — Pirart É. *L'Éloge mazdéen de l'ivresse : Édition, traduction et commentaire du Hōm Stōd*. Paris: L'Harmattan, 2004.

Probert 2015 — Probert Ph. *Early Greek relative clauses*. Oxford: Oxford Univ. Press, 2015.

Reichelt 1909 — Reichelt H. *Awestisches Elementarbuch*. Heidelberg: Winter, 1909.

Ross 1991 — Ross M. D. Refining Guy's sociolinguistic types of language change. *Diachronica*, 2001, 8/1: 119–129.

Rubinchik 2001 — Рубинчик Ю. А. *Грамматика современного персидского литературного языка*. М.: Восточная лит-ра, 2001. [Rubinchik Yu. A. *Grammatika sovremennogo persidskogo literaturnogo yazyka* [Grammar of Modern Standard Persian]. Moscow: Vostochnaya Literatura, 2001.]

Salvini 2008 — Salvini M. *Corpus dei Testi Urartei*. Vol. 1: *Le iscrizioni su pietra e roccia*. I : *Testi*. Rome: Centro Nazionale delle Ricerche, 2008.

Schmitt 1977 — Schmitt R. Königtum im Alten Iran. *Saeculum*, 1977, 28/4: 384–395.

Schmitt 1999 — Schmitt R. *Beiträge zu altpersischen Inschriften*. Wiesbaden: Reichert, 1999.

Schmitt 2007 — Schmitt R. *Grammatik des klassischen Armenischen*. 2. Auflage. Innsbruck: Institut für Sprachen und Literaturen der Universität Innsbruck, 2007.

Schmitt 2009 — Schmitt R. *Die altpersischen Inschriften der Achaimeniden*. Wiesbaden: Reichert, 2009.

Sims-Williams 1989 — Sims-Williams N. Eastern Middle Iranian. *Compendium linguarum iranicarum*. Schmitt R. (ed.). Wiesbaden: Reichert, 1989, 165–172.

Sims-Williams 2000 — Sims-Williams N. *Bactrian documents from Northern Afghanistan I: Legal and economic documents*. London: Nour Foundation, 2000.

Sims-Williams 2007 — Sims-Williams N. *Bactrian documents from Northern Afghanistan II: Letters and Buddhist texts*. London: Nour Foundation, 2007.

Sims-Williams 2009 — Sims-Williams N. The Bactrian fragment in Manichaean script (M 1224). *Literarische Stoffe und ihre Gestaltung in mitteliranischer Zeit*. Durkin-Meistererernst D.et al. (eds.). Wiesbaden: Reichert, 2009, 245–260.

Skjærvø 2009 — Skjærvø P. O. Old Iranian. *The Iranian languages*. Windfuhr G. (ed.). London: Routledge, 2009, 43–195.

Steblin-Kamenskii 1987 — Стеблин-Каменский И. М. Ваханский язык. *Основы иранского языкознания. Новоиранские языки: восточная группа*. Расторгуева В. С. (отв. ред.). М.: Наука, 1987, 408–473. [Steblin-Kamenskii I. M. Wakhi. *Osnovy iranskogo yazykoznaniya. Novoiranskie yazyki: vostochnaya gruppa.* Rastorgueva V. S. (ed.). Moscow: Nauka, 1987, 408–473.]

Stilo 2005 — Stilo D. L. Iranian as buffer zone between the universal typologies of Turkic and Semitic. *Linguistic convergence and areal diffusion: Case studies from Iranian, Semitic and Turkic*. É. Á. Csató et al. (eds.). London: Routledge, 2005, 35–63.

Stolper 2004 — Stolper M. W. Elamite. *The Cambridge encyclopedia of world's ancient languages*. Woodard R. D. (ed.). Cambridge: Cambridge Univ. Press, 2004, 60–94.

Stolper, Tavernier 2007 — Stolper M. V., Tavernier J. An Old Persian administrative tablet from the Persepolis fortification. *Achaemenid Research on Texts and Archaeology*, 2007.001. URL: http://www.achemenet.com/pdf/arta/2007.001-Stolper-Tavernier.pdf.

Sundermann 1981 — Sundermann W. *Mitteliranische manichäische Texte kirchengeschichtlichen Inhalts*. Berlin: Akademie, 1981.

Tavernier 1999 — Tavernier J. Zum altpersischen *hy* /hya/. *Archív Orientální*, 1999, 67: 683–693.

Tavernier 2008 — Tavernier J. Multilingualism in the fortification and treasury archives. *Archive des fortifications de Persepolis : État de question et perspective de recherche*. Briant P. et al. (eds.). Paris: De Boccard, 2008, 59–86.

Tavernier 2018 — Tavernier J. Multilingualism in the Elamite Kingdoms and the Achaemenid Empire. *Studies in multilingualism, Lingua Franca and Lingua Sacra*. Braarvig J., Geller M. J. (eds.). Berlin: Max Planck Institute for the History of Science, 2018, 307–320.

Thomason 2001 — Thomason S. G. *Language contact: An introduction.* Washington, D.C.: Georgetown Univ. Press, 2001.

Tsabolov 1997 — Цаболов Р. Л. Курдский язык. *Основы иранского языкознания. Новоиранские языки II*. Ефимов В. А. (отв. ред.). М.: Восточная лит-ра, 1997, 6–96. [Tsabolov R. L. Kurdish. *Osnovy iranskogo yazykoznaniya. Novoiranskie yazyki II.* Efimov V. A. (ed.). Moscow: Vostochnaya Literatura, 1997, 6–96.]

Tsabolov 2001–2010 — Цаболов Р. Л. *Этимологический словарь курдского языка*. М.: Восточная лит-ра, т. I: 2001, т. II: 2010. [Tsabolov R. L. *Etimologicheskii slovar' kurdskogo yazyka* [Etymological dictionary of Kurdish]. Moscow: Vostochnaya Literatura, vol. 1: 2001, vol. 2: 2010.]

Utas 2013 — Utas B. Semitic in Iranian: Written, read and spoken language. *Linguistic convergence and areal diffusion: Case studies from Iranian, Semitic and Turkic*. Csató É. Á. et al. (eds.). London: Routledge, 2005, 65–77.

de Vaan 2003 — de Vaan M. *The Avestan Vowels*. Amsterdam: Rodopi: 2003.

de Vaan 2018 — de Vaan M. The noncanonical use of instrumental plurals in Young Avestan. *Farnah: Indo-Iranian and Indo-European studies in honor of Sasha Lubotsky*. Van Beek L. et al. (eds.). Ann Arbor: Beech Stave, 2018, 21–36.

Viti 2015 — Viti C. *Variation und Wandel in der Syntax der alten indogermanischen Sprachen*. Tübingen: Narr, 2015.

West 2011 — West M. L. *Old Avestan syntax and stylistics*. Berlin: De Gruyter, 2011.

Windfuhr 2009 — Windfuhr G. Dialectology and topics. *The Iranian languages.* Windfuhr G. (ed.). London: Routledge, 2009, 1–42.

Yakubovich 2002 — Yakubovich I. Mugh 1.I. revisited. *Studia Iranica*, 2002, 31/2: 231–253.

Yakubovich 2010a — Yakubovich I. *Sociolinguistics of the Luvian language*. Leiden: Brill, 2010.

Yakubovich 2010b — Yakubovich I. Morphological negation in Urartian. *Aramazd (Armenian Journal of Near Eastern Studies)*, 2010, 5/1: 141–165.

Zólyomi 2011 — Zólyomi G. Akkadian and Sumerian language contact. *The Semitic languages: An international handbook*. Wenninger S. (ed.). Berlin: De Gruyter, 2011, 396–404.