



HAL
open science

Looking for the Middle Way: Voice and Transitivity in Complex Predicates in Iranian

Agnes Korn

► **To cite this version:**

Agnes Korn. Looking for the Middle Way: Voice and Transitivity in Complex Predicates in Iranian . *Lingua*, 2013, Special Issue: Complex Predicates, 135, pp. 30-55. <10.1016/j.lingua.2013.07.015>. <halshs-01340632>

HAL Id: halshs-01340632

<https://shs.hal.science/halshs-01340632v1>

Submitted on 1 Jul 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License

[NOTICE: this is the author's version of a work that was accepted for publication in *Lingua*. Changes resulting from the publishing process, such as editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document.

A definitive version was subsequently published in *Lingua* 135, pp. 30-55, <http://dx.doi.org/10.1016/j.lingua.2013.07.015>

**Looking for the Middle Way:
Voice and Transitivity in Complex Predicates in Iranian ***

Agnes Korn

Abstract

This article explores the emergence of complex predicates in Persian with a focus on voice and transitivity. It argues that the rise of CPs is linked to the development of the verb pair “do” and “become”, which encode the features called Instigation [+INST] and Affectedness [+AFF], respectively, by Næss 2007. While these features are characteristic for prototypical agents and patients, respectively, taken alone they are more general than that, making the two verbs “underspecified”, a typical characteristic of light verbs as noted by Megerdooomian 2012 and others. The distribution of the verbs “do” and “become” is parallel to the domains of the Old Iranian active and middle (mediopassive) voice; it thus mirrors the voice opposition whose morphological marking is lost within Middle Iranian. With “do” and “become” as its centre, the system integrates additional verbs such as “hold” and “give; put” on the [+INST] side and verbs of movement on the [+AFF] one. The same verbs are also used as auxiliaries for periphrastic formations such as the potential construction, the transitive preterite and the analytic passive, suggesting that grammaticalisation of auxiliaries and the development of light verbs are parallel processes the precise similarities and differences of which remain to be investigated. Here as elsewhere, the somewhat fragmentary evidence of early stages of Iranian is supplemented by data from languages that have found themselves under Iranian influence, providing details which are crucial to complete the picture.

Keywords:

Iranian languages, Persian, historical linguistics, complex predicates, voice, middle, transitivity, denominatives

* Abbreviations: ACC accusative; AFF affectedness; AOR aorist; CAUS causative; COND conditional; COP copula; CP complex predicate; DAT dative; DEM demonstrative pronoun; DIR direct case; F feminine; FUT future; GEN genitive; IE Indo-European; INF infinitive; INJ injunctive mood; INST instigation; IPF imperfect; IPFV imperfective aspect; IPR imperative; Ir. Iranian; ITR intransitive; LOC locative; M masculine; MDL middle; N neutre; NEG negation; NOM nominative; OBJ object case; OBL oblique case; OPT optative; PASS passive; Persian = Standard Modern Persian (unless otherwise noted); PF perfect; PFP perfect participle; PFV perfective aspect; PL plural; PN name; POT potential; PP past participle; PRO pronominal clitic; PRS present (stem); PST past (stem); PTC particle (in some cases: uninflected former participle); REL relative pronoun; SBJ subjunctive; SG singular; SUB subordinator; TR transitive; VOL volition. Nominals not marked for number are in the singular, those not marked for case are in the nominative or direct case, Old Ir. verb forms not marked for voice are in the active. In some instances, forms which are at the same time e.g. NOM and ACC are glossed only for the form that is actually relevant in the given clause.

Examples are not always entirely identical to the version in the publications quoted, as morpheme separating hyphens and glosses have been added (or, in some case, modified) and some translations are more literal than in the original; the transcription of some examples is also modified in some cases and added in some others. For examples from Old Persian and Middle Iranian, the first line gives the transliteration, the second the transcription.

Quotes from works written in other languages than English are my translations.

1. Introductory summary

[p. 31 of the printed version]

In Persian as well as in many other contemporary Ir. languages complex predicates often come in pairs, with one member functioning as active or transitive member (in Persian with *kardan* “to do”, *zadan* “to hit / affect”, etc.), the other one as its passive or intransitive counterpart (with *šodan* “to be, become”, *x^wordan* “to eat / hit”, etc.), as in (1).

1) New Persian

1a) ¹	ACTIVE:	<i>tarjome</i>	<i>kardan</i>	PASSIVE:	<i>tarjome</i>	<i>šodan</i>
		translation	do.INF		translation	become.INF
		"to translate"			"to be translated"	
1b)	TR:	<i>penhān</i>	<i>kardan</i>	ITR:	<i>penhān</i>	<i>šodan</i>
		hiding	do.INF		hiding	become.INF
		"to hide (TR)"			"to hide (ITR)"	

While many studies have looked at the matter from a theoretical perspective, studies of its historical development are rare. This paper thus proposes to look at possible origins and logics of the distribution of light verbs.

I will argue that pairs of complex predicates such as (1) originally encode the opposition of the features [+INST] (instigation) and [+AFF] (affectedness), as defined by Næss 2007. The most wide-spread light verbs are “do” (symbolized with *KAR* in this paper to indicate the Old Ir. root most commonly used for the purpose in Iranian) vs. “be, become” (here noted as *BAW*). That these two items should be the basis of the system appears to be confirmed by a large number of Ir. languages showing “do” vs. “become” as default light verbs, including several which use etymologically unrelated material to indicate the same opposition.

Crucially, the features relevant here do not constitute a binary opposition, i.e. *KAR* vs. *BAW* do not encode e.g. [+/-INST] nor [+/-AFF]). I suggest that this asymmetry historically lies at the heart of the system, and that it is a reason for the flexibility of the system, permitting the integration of additional light verbs.

Looking out for the history of the phenomenon, it springs to mind that Old Iranian does not show light verbs corresponding to the modern ones, but *KAR* and *BAW* are used in a distribution matching the opposition in (1) in the so-called potential construction (“be able to do X”), which is composed of the verbal adjective (later on the past stem) of the main verb plus a finite form of *KAR* “do”. This pattern is attested in Old Persian and in several Middle and New Ir. languages. It has a counterpart with *BAW* “be, become”, which covers not only the passive (“X can be done”), but also verbs that tend to feature the middle voice cross-linguistically, viz. “wear”, “lie”, etc. (thus e.g. “be able to lie down”). Subjects of middle verbs and of passives share the feature [+AFF], so that *BAW* encodes [+AFF] in the potential construction.

¹ Examples from Windfuhr (1979:113). For further such examples, see e.g. Ahadi (2001:271f.).

I argue that it is no coincidence that passives are grouped together with middles in the potential construction; this covers the range of the so-called mediopassive of ancient Indo-European languages. *BAW* thus becomes an analytical or periphrastic way of expressing the inherited mediopassive, which disappears in early Middle Iranian.

While the potential construction may have provided an entry point for *BAW* into periphrastic constructions, Semitic evidence suggests that *KAR* was the first verb to emerge as a light verb. Aramaic data indicate that (some variety of) Old Persian showed combinations of *KAR* with nouns in proto-complex predicate patterns. Syriac data suggest the same for Middle Persian, the pattern being apparently somewhat more common at this stage.

KAR expressing [+INST], a causative semantics in the broad sense, recalls the fact that it turns up in complex predicates besides synthetic causative verbs in Middle Persian, thus e.g. (2):

2) Middle Persian²

2a) *nām* *kardan* / *nām-ēn-īdan* / *nāmīdan*
 name do.INF name-CAUS/DENOM-INF name-INF
 “to name, to call”

2b) *handarz* *kardan* / *handarz-ēn-īdan*
 advice do.INF advice-CAUS/DENOM-INF
 “to give advice”

The Middle Persian suffix *-ēn-* is used for both causatives and denominatives; *KAR* as a light verb may thus have emerged as a periphrastic way of expressing causatives and also denominatives. The facts that *-ēn-* is likely to be a [p. 32] Middle Iranian formation and that CPs appear to oust Middle Persian causatives / denominatives in the first place would explain why combinations of *KAR* with noun, and by extension, the system of complex predicates as a whole, did not become prominent earlier.

Armenian furnishes further data potentially relevant for the history of complex predicates in Iranian: in addition to “do” and “be”, Armenian uses the verbs “hit”, “give” and “hold/have” paralleling the use of Persian *zadan*, *dādan* and *dāštan*. Numerous Armenian complex predicates are obvious calques on the Iranian model, many indeed using an Iranian noun. While it is not clear at which stage, and from which Western Ir. language, Armenian adopted the pattern, this particular selection of light verbs appears to point to the existence of a stage with several verbs for the [+INST] slot (“do”, “hit”, “give”, “hold/have”) while “be” was the only light verb to encode [+AFF]. This might be further evidence for the primary role of the [+INST] pattern in CPs, and for *KAR* “do” being established as the first light verb.

The use of verbs of movement as light verbs marking [+AFF] is likely to be a subsequent addition to the system. Verbs such as “come”, “move forward”, etc. are at the same time employed for the newly emerging analytical passives that replace earlier synthetic

² (2a) with *nāmīdan* from Sundermann (1989:152), the others from Ciancaglini (2011:11f.).

constructions. Periphrastic passives being grouped with [+AFF] complex predicates appears to underline the importance of the feature [+AFF] for the constitution of the system of complex predicates in the history of Iranian. At the same time, this line of argument suggests a parallelism in the development of complex predicates and periphrastic constructions, thus similarities in the development of light verbs and auxiliaries, with possible implications on the ongoing discussion about pathways of grammaticalisation.

Before discussing these issues broadly in the sequence outlined above, Section 2. will present some more general points. Section 5. summarises the findings and surveys the preceding ones with regard to inhowfar they confirm or disprove other authors' arguments about the development of complex predicates cross-linguistically.

2. Topics and framework

2.1 The phenomenon to be discussed

The term “complex predicate” (CP) as discussed in this paper refers to combinations of a noun (or another non-verbal element) with a so-called light verb, thus to the structure termed “light verb construction” by Bower (2008:163).³ In instances such as (1), the verb is semantically bleached (or “underspecified”, Megerdooian 2012:180). The same verb may have its full semantic force in other contexts and is then a “heavy verb”.⁴ Light verbs are commonly a “small closed class” while the elements that may be combined with it to give a CP are “an open class”.⁵

1') New Persian

1a)	ACTIVE:	<i>tarjome</i>	<i>kardan</i>	PASSIVE:	<i>tarjome</i>	<i>šodan</i>
		translation	do.INF		translation	become.INF
		"to translate"			"to be translated"	
1b)	TR:	<i>penhān</i>	<i>kardan</i>	ITR:	<i>penhān</i>	<i>šodan</i>
		hiding	do.INF		hiding	become.INF
		"to hide (TR)"			"to hide (ITR)"	

³ Terminology has not been uniform. Patterns called CPs here are termed “conjunct verbs” by Bashir (2008:65), following usage in descriptions of New Indic languages, while “compound verbs” usually refers to combinations of a verb stem / root (main verb) and a finite light verb in this tradition (cf. Schmidt 2003:337-339). Various authors subsume such verb + verb combinations under the term “complex predicate” as they describe a single event, and distinguish them from serial verbs, which describe more than one event (whether the latter is indeed so is not uncontroversial either, see Foley 2010). Some other authors include serial verbs within their definition of “complex predicate”, defining the term as patterns that are monoclausal syntactically (cf. Baker / Harvey 2010:13). Monoclausality differentiates CPs from control constructions (Butt 2010:49, 57-59). The term “light verb” appears to go back to Jespersen (1965/VI:117), cf. Butt (2010:48).

⁴ Form identity to a full verb is a “central characteristic of light verbs” according to Butt (2010:53); cf. Section 5.4 for further discussion.

⁵ Thus Baker / Harvey (2010:15) for constructions in Marra (Australia) that would be called “compound verbs” by the convention in note 3.

Light verbs host the inflection of the verbal phrase, “carrying all the syntactic features generally attached to the verb: tense, mood, negativity, person, and number” (Sheintuch 1976:137).⁶ In addition, light verbs also convey properties of the CP such as transitivity, voice and *aktionsart*. This suggests that light verbs are a periphrastic expression of these categories, and equivalent to derivational morphology in other languages (Windfuhr 1979:117, quoting Telegdi 1951; Butt 2010:52), just as in (2a) *nām kardan* “to name, call” is a periphrastic way of saying the same thing as *nām(ēn)īdan*. [p. 33] This feature also makes CP patterns a cross-linguistically very common device for incorporating loans (see e.g. Baker / Harvey 2010:16 for Marra, Australia, Fritz 2009:16, 49f., quoting Fragner 1999:27-29 for various languages of Eurasia).

As noted by many authors, CPs are a pervasive feature of New Persian to an extent that the number of simple verbs is rather small, and many concepts for which in other languages a simple verb would be used are expressed by a complex predicate. Thus a Persian CP “is equivalent semantically and structurally” (Megerdooian 2012:190) to a simple verb. This is an important difference between the status of Persian CPs and apparently parallel constructions seen in European languages such as Hungarian, German or French or other languages: as pointed out by Telegdi (1951:317f., 324), the German and Persian CPs meaning “to answer” (3) are only superficially equivalent; the crucial difference is that German *Antwort geben* coexists with the simple verb *antworten* while Persian CPs are characterised by the absence of a corresponding simple verb.

3)	German	<i>Antwort geben</i>	/	<i>antworten</i>
	Persian	<i>javāb dādan</i>	/	∅
		answer give.INF		answer.INF
		“to answer”		

In languages like German, CPs are thus periphrastic expressions in the true sense, “an indirect way of expressing oneself, a detour” (Telegdi *ibid.*), and situated on the level of the lexicon while CPs in Persian are a grammatical phenomenon.⁷ So the difference does not concern the frequency of CPs, but their status in the grammar of the language. A discussion of how complex predicates may have emerged in Persian and other Iranian languages thus does not only need to account for the mechanism motivating patterns such as those in (1), but also of how these gained their status here symbolised by the contrast in (3).

The opposition of light verbs in examples such as (1a) has frequently, and as early as the 17th century (cf. Windfuhr 1979:117), been described as contrasting actives (with *kardan* “to do” or *zadan* “to hit / affect”, etc.) to passives (with *šodan* “to become”, *x^wordan* “to eat / hit”, etc.). Indeed the marking of transitivity by the choice of the light verbs appears to be rather

⁶ Butt (2010:52) considers inflection and conveying tense / aspect information “not a typical characteristic of light verbs” cross-linguistically.

⁷ This position agrees with the majority view in approaches such as the one by Folli et al. 2005.

common cross-linguistically.⁸ However, it has also been noted that Persian CPs are by no means a uniform category. Pairs such as (1b) rather oppose transitives to intransitives (cf. the remarks in passing by Windfuhr 1979:118), with the same light verb as in (1a). But this opposition still falls short of accounting for the distribution of light verbs, as these can convey a range of other properties as well, including *aktionsart* and other nuances.⁹ The present paper thus proposes that these issues should be discussed together with the various light verbs actually employed.

This approach naturally implies that the argument to be presented here shares the view that “light verbs are not in fact semantically empty” (Megerdooian 2001: 100, similarly Folli et al. 2005, Butt 2010: 49-51 and many others), and that both the light verb and the non-verbal element contribute to the properties of the complex predicate.

2.2 The approach of this paper

The various features of Persian CPs, including their marking of voice, transitivity, *aktionsart*, etc., have widely been discussed from descriptive and theoretical viewpoints. Much of the discussion has focussed on the respective contribution of the light verb and the preverbal element to the features of the complex predicate.¹⁰

From a historical perspective, the fact that light verbs “neither retain their full semantic predicational content, nor are they semantically completely empty” (Butt 2010:48) might agree with the general assumption that light verbs develop from heavy verbs, but it then would remain to be shown “which parts of the predication are supposed to have been lost as part of the historical change” (Butt 2010:49). While Butt goes on to say that “there is no documented evidence of such a historical change”, I will try to present some data that may serve to illustrate such a change.

In so doing, I will make use of the descriptive instrumentarium used by Næss 2007 in her discussion of transitivity features. She describes the participants of an event by three features: volition [VOL] characterises agents, “volitional undergoers” and “frustratives” (i.e. “a participant who is volitionally involved in the act but does not actually instigate anything”, Næss 2007:99), and excludes involuntary actors; instigation [INST] is a property of agent, force and instrument [p. 34] (to the exclusion of volitional undergoers); and affectedness [AFF] applies to instrument, volitional undergoers and patients. The agent of a prototypical transitive sentence shows the properties volitionality and instigation, but not affectedness (i.e.: [+VOL, +INST, -AFF]) while the prototypical patient has the opposite characteristics ([-VOL, -INST, +AFF]). Næss thus predicts, among other things, that an agent whose

⁸ Cf. e.g. the data from Amharic and other Ethiopic languages presented by Amberber (2010:255, 299-303) and the Australian data in Baker / Harvey (2010:14-16).

⁹ Some examples are given in Ahadi (2001:262f.). For examples of certain Persian light verbs used in transitive as well as in intransitive constructions, see e.g. Dabir Moghaddam (2006:84-88).

¹⁰ Thus, aspectual properties have been assumed to be contributed by the light verb (Karimi-Doostan 1997 and 2005, who includes telicity here, and Megerdooian 2001) or by the preverbal element (thus Folli et al. 2005) or both (Pantcheva 2009).

features differ from the prototypical constellation (e.g. by being [+VOL, +INST], but also [+AFF], as in “eat”) will be marked differently from a prototypical one in some languages.¹¹

I will also make use of the discussion by Bower 2008, who offers a diachronic survey of complex predication (in a larger sense) and of her definition (Bower 2008:163) of light verbs as having “an empty (...) or reduced (...) argument structure”; as not assigning “theta roles (that is, they do not assign a semantic role to one of more of their arguments)”;¹² and as semantically deficient.

In what follows, the notations *KAR* and *BAW* will be used to refer to the verbs “do” and “become”, which in the various Ir. languages derive from the Proto-Iranian roots **kar* and **baw*, respectively (cognates of the Old Indic roots *kr̥* and *b^hū*), but to some extent also to etymologically unrelated verbs of the same meaning. Pink colour [bold, in the printed version] marks the multifunctional verbs under discussion, i.e. light verbs as well as light-verbs-to-be and also some auxiliaries under focus in this paper.

Introducing the Iranian languages discussed in this paper in a very rough and brief way, I list them broadly from east to west for each period, which at the same time arranges them in chronological order of their attestations so far as Old and Middle Iranian is concerned:

- Old Iranian (ca. 1000 to ca. 300 BC): Avestan and Old Persian;¹³
- Middle Iranian (ca. 300 BC to ca. AD 700): East Iranian: Khotanese (texts from ancient Turkestan, present Xinjiang, China; chiefly Buddhist texts), Sogdian (along the Silk Road, Central Asia into China; texts from various religions), Xwarezmian (grouped with Middle Iranian for reasons of its grammatical structure although most texts are from the Islamic period); West Iranian: Parthian (here quoted: Manichean religious texts), Middle Persian;
- contemporary Iranian languages (since the advent of Islam): East Iranian: Ossetic (Caucasus), Yaghnobi (a small minority language of Tajikistan, also noteworthy for sharing a number of isoglosses with Sogdian), Wakhi (a minority language along the upper Amu Darya and in northern Pakistan), Pashto (Afghanistan, Pakistan); West Iranian: Zazaki (spoken in eastern Anatolia in close contact with Kurmanji), Balochi (now in southeast Iran, southwest Afghanistan, western Pakistan), Gilaki (Iran, province of Gilan south of the Caspian Sea), Kurmanji (“Northern Kurdish”; Anatolia, Northern Iraq, etc.), (New) Persian.
- Data from contact languages is important to complement the sometimes fragmentary attestation of earlier periods of Iranian. Languages relevant for the present purposes include the Elamite and Babylonian versions of the Old Persian inscriptions, Aramaic as administrative language of the Achaemenid empire, and Armenian for the Middle Iranian period.

¹¹ Cf. specifically Næss (2007:44) and her examples pp. 116f. for further details. Note that Næss’ “prototypical” does not imply unmarkedness nor high frequency. Næss’ approach has some similarities to that applied by Pantcheva 2009, who distinguishes the “subevents” initiation (subject: Initiator); process (subject: Undergoer); and result (subject: Resultee).

¹² Note, however, that Bower (2008:166f. n. 8) quotes, apparently approvingly, Butt’s differentiating of light verbs from auxiliaries by (inter alia) the criterion that light verbs “assign theta roles and interact with argument structure”.

¹³ The Old Persian examples quoted below are all from the inscriptions of Darius I (ruled 522-486 BC).

3. The potential construction and the middle voice

As I will argue in Section 4., the Old and Middle Iranian evidence for complex predicates is somewhat inconclusive. Additional material appears to be necessary to account for the emergence of complex predicates as a system. This section will thus look at a periphrastic verbal pattern that could have provided a entry point for verbs that later on become light verbs. This pattern, the so-called potential construction, at the same time highlights an issue that appears important also for the development of complex predicates, viz. the distribution of the auxiliaries *KAR* and *BAW*.

3.1 The potential construction: Data

3.1.1 The pattern with *KAR*

While ancient Indo-European (IE) languages possess a large number of morphological means to express tense, mood and aspect categories in synthetic ways, there are also a few periphrastic verbal constructions. The potential construction is among the oldest patterns of verbal periphrasis in Iranian, and is found in numerous Middle and New Ir. languages.¹⁴ [p. 35] It is composed of a finite form of *KAR* “do” plus the past participle (which in later stages becomes the past stem) of the main verb. It can be used with intransitives (see below) and transitives, meaning “to be able to...”. The subject is in the nominative (Old Persian and most of East Iranian) or direct case. Example (4), from Parthian (Western Middle Iranian), illustrates the pattern, showing *windād* (the past stem of *wind-* “find”), combined with *kar-* “do” to yield “be able to find”.¹⁵

4) Parthian (Sundermann 1981:110, quoted from Durkin-Meisterernst 2002:58)

[...] ³ c	ʿy(myn	pnj)	[ʿyr](ʿn)	kd(ʿc)	...	(wy)nd ² d	krynd
aʃ	imīn	panʃ	īrān	kaḏāʃ	...	wind-ād	kar-ēnd
from	DEM.OBL.PL	five	thing.OBL.PL	ever		find-PST	do.PRS-3PL

“they can find [nothing of?] these five things”

According to Benveniste 1954, the construction originally means “succeed, achieve, complete an action”, from where the notions of both potentiality (chiefly in main clauses) and anteriority (mainly in subordinate clauses) are derived.¹⁶ The comparatively clearest Old Iranian candidate for a potential interpretation is the Old Persian sentence in (5).¹⁷ While the degree to which this construction was already grammaticalised in Old Persian is not entirely clear, the Elamite version available for some Old Persian inscriptions confirms the potential

¹⁴ See Gershevitch (1954:131), Zarshenas 2002 and Sims-Williams 2007 (with references to earlier discussions) for details on the potential construction, as well as the sources quoted below.

¹⁵ For the Parthian potential construction, see Sundermann (1981:59).

¹⁶ Emmerick (1987:284f.) notes this function also for main clauses in Khotanese.

¹⁷ Thus Reichelt (1931:257), Schmitt (1991:52, 2009:44), Sims-Williams (2007:381). Other examples interpreted as potential by Benveniste (1954:64) are seen as anterior by Sims-Williams (ibid.). See also note 24.

meaning of the construction.¹⁸

5) Old Persian (Dareios' inscription at Bisotun I. 48-50)¹⁹

<i>n-i-y</i>	<i>a-h</i>	<i>m-r-t-i-y (...)</i>	<i>h-y</i>	
<i>naiy</i>	<i>āha</i>	<i>martiya (...)</i>	<i>hya</i>	
NEG	be.PF.3SG	man.NOM.SG	REL.NOM.SG.M	
<i>g-u-m-a-t-m</i>	<i>x-š-ç-m</i>	<i>d'-i-t-m</i>	<i>c-x-r-i-y-a</i>	
<i>gaumātam</i>	<i>xšaçam</i>	<i>dītam</i>	<i>čaxr-iyā</i>	
PN.ACC.SG	lordship.NOM/ACC.SG	tear.PP.NOM/ACC.N.SG	do.PF-OPT.3SG	

“there was no man who could have torn the lordship off Gaumata”

The past participle agrees with the object, thus *dītam* “torn” with *xšaçam* “lordship”. The form in *-tam*, being the NOM/ACC.N.SG and the ACC.M.SG of the past participle, will surely have featured in the majority of instances. This is reflected by the morphology of the potential construction in archaic Middle Iranian: by the Sogdian preference for *-tu* (the regular outcome of **-tam*), as in (6), where a form going back to **krtam* “done” is combined with the finite verb *wanān* “I should do” (subjunctive 1SG),²⁰ and by Khotanese, where a form likewise deriving from Old Ir. **-tam* (*jsīdu* “deceived” in (7)) is used irrespective of the gender or number of subject and object.²¹

6) Sogdian (Yoshida 2009:302)

<i>ḳrtw</i>	<i>wnʹn</i>	
<i>əktu</i>	<i>wan-ān</i>	
do.PP.POT	do.PRS-SBJ.1SG	

“I should be able to do”

7) Khotanese (Emmerick 1987:279)

<i>ne</i>	<i>balysu</i>	<i>...</i>	<i>jsīdu</i>	<i>yan-īndā</i>	
NEG	Buddha.ACC		deceive.PP.ACC	do.PRS-3PL	

“they cannot deceive the Buddha”

¹⁸ Sims-Williams (2007:381f.). This also applies to the intransitive/passive counterpart with *BAW* (on which see below).

¹⁹ Reichelt (1931:257), Schmitt (1991:29, 52, 2009:44).

²⁰ Sims-Williams (2007:380, 382). The form in *-tu* is the ACC.SG.M (or NOM/ACC.SG.N) of the so-called light stems, which preserve more of the Old Ir. inflectional patterns (heavy stems have *-t* without ending). The form of the Sogdian past participle used elsewhere (e.g. in the passive) shows different morphology (suffix M *-tē*, F *-tča*, PL *-tēt*, Sims-Williams 2007:377). Sogdian varieties include Buddhist, Manichean and Christian Sogdian, etc.; these differ from each other in script and other details, causing a certain amount of variation also in the transcribed form.

²¹ There is agreement with the subject in the intransitive construction (cf. (13b) below, cf. Emmerick 1987:279f.). Khotanese also has a verb “be able”, disproving Benveniste’s (1954) assumption that the potential construction makes up for the absence of a verb in this meaning (Emmerick *ibid.*).

[p. 36] In Xwarezmian, the potential is apparently only employed negatively (“to be unable to ...”, cf. Sims-Williams 2007:378 n. 3). *KAR* has developed into a particle (=k-) that is suffixed to the past stem of the main verb, such as *pard* “restrain” in (8), and carries the verbal inflection.

- 8) Xwarezmian (MacKenzie 1971:40, Durkin-Meisterernst 2009:364)

kf^om^ony *prdkⁱ*
ka=fa=ma *ne=* *pard=k-i*
 for=PTC=PRO.1SG NEG= restrain.PST=POT.PTC-2SG
 “for you cannot restrain me”

The existence of the construction has been noted for a number of contemporary languages, but is usually not quite well documented. It may thus not be superfluous to present some instances from Balochi as an example for the potential in New Iranian. In (9) the past stems *sist* “broken” and *dāt* “given” are combined with a finite form of *KAR* to express “I am able to break (i.e. pick) / give”.

- 9) Balochi (Western group, Pakistan; Barker / Mengal 1969/I:344)

aga *taw* *drust* *nīwag-ā* *sist* *bi-kan-ay*
 if you all fruit-OBL break.PST SBJ-do.PRS-2SG
guṛā *man* *tarā* *yakk bār=ē* *dāt* *kan-īn*
 then I you.OBJ one share=one give.PST do.PRS-1SG
 “if you can pick all the fruit, then I will be able to give you one share”

Instead of the past stem used in (9) and (11) some dialects (also) use an enlarged form of the past stem, viz. the perfect participle in *-a(g)*, thus *āt-ag* “come” in (10).²²

- 10) Balochi (Western group, Turkmenistan; Axenov 2006:223)

man *bi* *tīā* *āt-ag* *na-kurt-un*
 I to you.LOC come.PST-PFP NEG-do.PST-1SG
 “I could not come to you”

Balochi dialects with ergative structures show ergative case marking in the past domain, as in (11), the agent being in the oblique case on account of “do” being a transitive verb.

- 11) Balochi (Pakistan)

11a) Southern *āyā* *kapt* *kut-ag=ā*
 DEM.PL.OBL fall.PST do.PST-PFP=COP.PST
 “they could have fallen” (Farrell 1990:75)

²² Instead of the past stem, the infinitive (= past stem + *-īn* or present stem + *-ag*, depending on the dialect) is also found (see Jahani 1999:128 for the active construction in Western and Southern Balochi). The Balochi potential was already noted by Mockler (1877:61, 71).

- 11b) Eastern *āhīā* *k^huθ* *na-k^huθ-a*
 DEM.OBL do.PST NEG-do.PST-PFP
 “s/he could not do it” (Gilbertson 1923:132)

3.1.2 The pattern with *BAW*

Several Ir. languages also show an intransitive and passive potential construction which, instead of *KAR*, combines the perfect participle / past stem with a finite form of *BAW* “be, become”, again with the subject in the nominative or direct case.²³ The passive means “something can / cannot be done” while the intransitive use is found with verbs such as “lie [p. 37] down” (12b), etc., or some verbs of movement (13b). The pattern may only accidentally be unattested in Old Persian, as the passive with *BAW* in anteriority function is found.²⁴

12) Sogdian²⁵

- 12a) TR *βzyz* *L^o* *βrt* *wn-³y=k²m*
 bad NEG carry.PP.POT do.PRS-2SG=FUT
 “you will not be able to bear the hardship”
- 12b) ITR *χw* *L^o* *np³st^o* *βw-t*
 DEM.NOM.SG.M NEG lie down.PP.POT become.PRS-3SG
 “he cannot lie down”
- 12c) PASS *wyt^o* *wβ²-t*
 say.PP.POT become.PRS.SBJ-3SG
 “[it] could be said”

13) Khotanese²⁶

- 13a) TR *biśśā* *ttātā* *ggurvī²ca* *haṃk^haštu* *yan-īndā*
 all.NOM/ACC.PL DEM.PL particle.NOM/ACC.PL count.PP.ACC do.PRS-3PL
 “they can count all these particles”
- 13b) ITR *ni* *biśśā* *hā* *tranda* *hām-āre*
 NEG all.NOM/ACC.PL yonder enter.PP.NOM/ACC.PL become.PRS-
 MDL.3PL
 “they cannot all enter”

²³ See Yoshida (2009:302) for Sogdian, Emmerick (1987:279-288; 2009:396f.; note that in Emmerick 2009, *ā* is frequently misprinted as *ā*, thus the examples need to be used with caution) for Khotanese and Sims-Williams 2007 for both (p. 382ff. on the possible origin of the form in *-ta* found in the Sogdian construction). In Khotanese, some intransitives unexpectedly use *yan-* “do”; these are verbs that also show pattern transitively in the formation of their perfect; conversely, some actives are found with *hām-* “become” (Emmerick 1987:287f.). Durkin-Meisterernst (2002:58f.) discusses two possible examples from Parthian, but prefers alternative interpretations for both.

²⁴ Cf. Sims-Williams (2007:381). See note 17 for the anteriority function.

²⁵ (12a) and b) from Sims-Williams (2007:378, Buddhist Sogdian), c) from Gershevitch (1954:131, Manichean Sogdian).

²⁶ (13a) and c) from Emmerick (1987:280f.), b) from Sims-Williams (2007:380). For more examples of the Khotanese potential construction, see Emmerick (1968:111, 152).

13c)	PASS	<i>puñīnai</i>	<i>hambīsä</i>	<i>ne haṃk^häštä</i>	<i>häm-äte</i>
		of merit	heap.NOM.SG	NEG count.PP.NOM.SG	become.PRS-MDL.3SG

“the heap of merits cannot be counted”

Although rarely mentioned, the passive / intransitive construction also exists in some contemporary Ir. languages. For Yaghnobi, Klimčickij (1937:15-18) notes the use of *BAW* for passives and intransitives while Xromov (1972:44f.) adds that *BAW* is also used “for transitives used without logical subject for the action”. The pattern is only found negatively, and the construction expresses “the impossibility of the realisation of the action” (Xromov *ibid.*), as in (14).

14) Yaghnobi (Xromov 1972:45)

14a)	TR	<i>moy</i>	<i>na-žoyt</i>	<i>kun-im=išt</i>
		we	NEG-read.PST	do-1PL=PRS
		“we cannot read”		

14b)	ITR	<i>be</i>	<i>hamra</i>	<i>na-ed</i>	<i>višt</i>
		without	comrade	NEG-go.PST	become.2SG
		“one (lit.: you) can’t go without a comrade”			

In Balochi, some dialects employ *KAR* also for intransitives (10, 11a), but there is also a pattern with *BAW*, as in (some of?) Eastern Balochi (15).²⁷

15) Balochi (Eastern group, Pakistan; Gilbertson 1923:133)

15a)		<i>bīθ-a</i>	<i>bī-</i>
		become.PST-PFP	become.PRS
		“to be able to be / become”	

15b)	<i>ma yak</i>	<i>māh-a</i>	<i>angō</i>	<i>rasiθ-a</i>	<i>na-bā</i>
	I	one month-OBL	there	arrive.PST-PFP	NEG-become.PRS.1SG
	“I shall not be able to arrive there in a month”				

[p. 38] For Turkmenistan Balochi, Axenov (2006:224) notes the pattern under the title “impossibility construction”. Two of his examples express general impossibility (16), and the third one a more specific situation (17).²⁸

²⁷ The presence of this construction in more Balochi dialects than hitherto known is suggested by the fact that it appears to have been copied into Brahui (a Dravidian language that has been heavily influenced by Balochi). The Brahui construction combines an infinite form of the main verb with “do” for transitives and “become” for passives (cf. Barjasteh Delforooz 2007, Section 2.3, no intransitive examples given).

²⁸ Axenov notes that the pattern employs “inflected forms” of “become”. His examples all feature the 3SG indicative present, raising the question whether the pattern might be limited to this form.

16) Balochi (Western group, Turkmenistan; Axenov 2006:224)

16a) *rōč-ayā sar būt-ag =a na-bīt*
 sun-LOC head become.PST-PFP IPFV NEG-become.PRS.3SG
 “one cannot reach (lit.: to, as far as) the sun”

16b) *pa γulp čēr kurt-in sātīt-ag =a na-bīt*
 to lock under do.PST-INF keep.PST-PFP IPFV NEG-become.PRS.3SG
 “one cannot hide [a secret] under lock”

17) Balochi (Western group, Turkmenistan; Axenov 2006:224)

ēšī dast čōt mant
 DEM.GEN hand.DIR crooked remain.PST
rāst kurt-ag =a na-bīt
 right do.PST-PFP IPFV NEG-become.PRS.3SG
 “his hand remained crooked, it cannot be made straight”

In Pashto and some other contemporary Ir. languages, the use of *BAW* has been generalised to include also transitive verbs. It employs the conditional, which is formed by adding the suffix *-āy* to the past stem or the infinitive (Lorenz 1979:183f.) to give instances such as (18).²⁹

18) Pashto (Lorenz 1979:183)

18a) *darēd-əl-āy šw-ē*
 stop.PST-INF-COND become.PST-2SG
 “you could stop”

18b) *stāsī xabarī nə-š-əm pōhēd-əl-āy*
 your(PL).DIR word.DIR.PL NEG-become.PRS-1SG understand.PST-INF-COND
 “I cannot understand your words”

3.2 The development of the potential construction and the verbs it employs

Perhaps somewhat unexpectedly for ancient IE languages, periphrastic constructions begin to emerge already in Old Iranian, one of the oldest being the potential construction, which shows various changes in the ca. 2500 years of its observable history.

The first one involves the availability of alternative readings. Initially, the potential is interpretable as employing a full verb: concerning the pattern with *KAR*, Emmerick (1987:279) notes that (7) “means literally ‘They do not make the Buddha deceived.’”

7) Khotanese (Emmerick 1987:279)

ne balysu ... jsīdu yan-īndä
 NEG Buddha.ACC deceive.PP.ACC do.PRS-3PL
 “they cannot deceive the Buddha”

²⁹ For further Pashto examples, see Robson / Tegey (2009:751f.). Morgenstierne (1927:77) notes a similar use of “become” for some Western Ir. varieties, but the material would need to be checked.

Similarly, “the passive of the potential construction is formally identical with the ordinary periphrastic passive” in Khotanese (Emmerick 1987:281) and in Old Persian (Sims-Williams 2007:382). For instance, (13b) can also mean “the heap of merits is not / will not be counted”.³⁰

[p. 39] However, the potential construction develops into a morphologically separate pattern already in Khotanese, since the form of the ACC.SG.M or NOM/ACC.SG.N is generalised in the pattern with *KAR*, abandoning agreement of the perfect participle with the object.³¹ While (7) is still grammatical with the alternative translation “make the Buddha (ACC.SG) deceived (ACC.SG)”, such a reading is not available for other objects, so that e.g. (13a) cannot be read as *“(they make the particles (PL) counted (ACC.SG))”. As soon as the perfect participle then becomes the past stem, it cannot be interpreted as a predicative adjective anyway.

In Sogdian, there is a different morphology of the participle form employed for both the transitive and the intransitive pattern. Thus a passive potential such as (12c) is formally different from a periphrastic passive “it is said”;³² demonstrating that Benveniste 1954, who argues this “is an idiom which belongs in the lexicon rather than the grammar”, cannot be right for this stage of the construction (Sims-Williams 2007: 377f.).³³ The same applies to the Pashto formation, employing a derivative of the past stem for the potential.

Sogdian also shows another formal difference between the potential and the periphrastic passive since the negation precedes the finite verb in the passive, but is prefixed to the main verb in the potential construction.³⁴

Another morphosyntactic difference between *KAR* and *BAW* in the potential construction and tense or voice auxiliaries is the argument structure seen in the Balochi patterns. In compound tense forms such as the perfect or past perfect, the case of the subject is determined by the transitivity of the main verb while the intransitivity of the auxiliary (i.e. the copula or *BAW*) is irrelevant (e.g. *āyā* (DEM.OBL.PL) *gušta bīta* “they had said”, Farrell 1990:74). In the potential construction, on the other hand, the transitivity of the main verb is irrelevant, as shown by (11), where ITR *kapt* “fallen” is treated in the same way as TR *k^huθ* “done”, and “the construction is seen as a transitive one” due to the presence of the TR finite verb *KAR* (Farrell 1990:58).

³⁰ For this type of periphrastic passive, see Emmerick (1987:278, 281; 2009:397).

³¹ Emmerick (1987:279) notes the form as being NOM.ACC.N, but the ACC.M has the same form (Sims-Williams 2007:382 n. 17).

³² Sims-Williams (2007: 377); see note 20.

³³ In Balochi, the main verb is found in various derivatives of the past stem in the passive (cf. Jahani / Korn 2009:663), but from the available data, there does not seem to be a dialect using the potential with “become” in which the passive would have a different form of the main verb than in the potential.

³⁴ Cf. (12) and the negated examples in Gershevitch (1954:131f., 125f.), respectively. In Yagnobi (where the negation is likewise prefixed to the main verb in the potential), a periphrastic passive resembling the potential construction does not seem to be in use.

11') Balochi (Pakistan)

11a) Southern *āyā* *kapt* *kut-ag=ā*
DEM.PL.OBL fall.PST do.PST-PFP=COP.PST
“they could have fallen” (Farrell 1990:75)

11b) Eastern *āhīā* *k^huθ* *na-k^huθ-a*
DEM.OBL do.PST NEG-do.PST-PFP
“s/he could not do it” (Gilbertson 1923:132)

In the potential construction, *KAR* and *BAW* thus show properties different from those exhibited by tense or voice auxiliaries. They also differ from auxiliaries in the fact that they form a pair right from the start (while there is usually only one auxiliary each for the passive or the perfect).³⁵ Noteworthily, the verbs employed as “become” in the various Ir. languages are not etymologically identical: while many languages use *BAW*, Khotanese has *hām-* of unclear etymology³⁶ and Pashto uses *šw-/š-* (probably a cognate of New Persian *šodan*, etymologically “move forward”³⁷). The potential construction thus presents itself as not a fossilised pattern, but a construction that remains functional even when new etymological material moves in.

So far as the distribution of *KAR* vs. *BAW* is concerned, it springs to mind that intransitives such as “lie, sit” and verbs of movement such as “enter” (Sims-Williams 2007:380f.) employ *BAW*, suggesting that *BAW* is selected if the subject of the sentence shows the property labelled [+AFF] by Næss (see Section 2.2).³⁸ For Yaghnobi, Xromov translates the pattern with *BAW* in (14b) with *ne-l’zja* “it is not possible” while he renders negated examples with *KAR* with *ne možeš* “you cannot...” (thus also the other persons). This could indicate a general impossibility in the first pattern, thus rather a feature of the person affected (again [+AFF]) than of the action that should be carried out. Xromov also includes transitives “without logical subject” in the group that employ “become”, but does not give an example. At any rate, if there is no logical subject, the object must be the focal point, so perhaps the Balochi ex. (17) “his hand (...) cannot be made straight” could [p. 40] be an instance of what Xromov has in mind, showing *dast* “hand” as grammatical subject, but semantically a patient, and similarly for (16), expressing the general impossibility of certain acts.

BAW being also used for the passive potential groups intransitives together with passives against transitive actives. So the languages that use two different verbs in the potential construction select *BAW* if the subject of the sentence shows the property [+AFF], otherwise *KAR*. In a number of instances, but certainly not all of them, the subject of the

³⁵ However, see Section 5.3 for different perfect formations depending on transitivity.

³⁶ Emmerick (1968:152, 242).

³⁷ Morgenstierne (2003:78).

³⁸ However, the Sogdian infinite construction employing the verbal noun *kʷ* “deed, work” appears to be used for any verb, including “die” (cf. Sims-Williams 2007:381f.), apparently with a generalisation of the transitive pattern.

pattern with *BAW* also shows the feature [-INST], thus two of the features characteristic for Næss' prototypical patient. The grouping of [+AFF] intransitives with passives recalls the use of the middle voice in ancient IE languages, which is employed for the same two groups of patterns, and raises the hypothesis that *BAW* might be a periphrastic expression covering the slot of the middle voice when the latter disappears in the course of the development of the individual languages. The next section will thus present the middle voice in Old and Middle Iranian.

3.3 The middle voice: data

Ancient Indo-European languages have one set of verbal endings for the active voice and a second one for the middle. The IE middle has also been called “mediopassive” because it is not only used in functions such as found for middles cross-linguistically, but also for the passive.³⁹ Like other IE languages, Indo-Iranian also developed a separate passive voice.⁴⁰

Examples illustrating the functions of the Old Ir. middle in contrast to active uses of the same verb are presented in (19-22), taken from Sims-Williams 1994.⁴¹

19) Old Persian

19a) *kāram* *pā-dī* ACTIVE
 people.ACC.SG protect.AOR-IPR.2SG
 “protect the people (imperative)” (Dareios, Persepolis e l. 21)

19b) *patīpaya-uvā* REFLEXIVE
 protect.PRS-MDL.IPR.2SG
 “protect yourself” (Dareios, Bisotun IV. 38)

20) Avestan

20a) *hačáti* *dim* ACTIVE
 follow.PRS.3SG DEM.ACC
 “it follows him” (Yasht 19. 68)

20b) *nōiṭ daēnā* (...) *hačánte* RECIPROCAL
 not religion.NOM.PL follow.PRS.MDL.3PL
 “our religions do not follow (= conform to) each other” (Yasna 45. 2)

21) Avestan

21a) *yōi* ... *aēšəməm* *varədən* ACTIVE
 REL.PL.M wrath.ACC.SG grow.PRS.INJ.3PL
 “[those] who increase anger” (Yasna 49. 4)

³⁹ Cf. e.g. Fortson (2010:89f.).

⁴⁰ This passive in *-ya-* is somewhat peripheral to the verbal system in that it is based on the root, not on the stems (present, aorist, perfect), on which the remaining verb forms are based. The difference to the middle is that it can only be used for the passive while the middle is also used in several other functions.

⁴¹ For Avestan, see also Kellens (1984:39-80) and Reichelt (1909:296-299); cf. also Skjærvø (2009:139).

21b) *daŋhuš* ... *varədataē=ča* INTRANSITIVE
 country.NOM.SG grow.PRS.MDL.SBJ.3SG=and
 “and the country will increase” (Yasht 13. 68)

22) Avestan (Kellens 1984:46)

22a) *yō* *vō* *hubərəta* *barāti* ACTIVE
 REL.NOM.SG.M you.PL.ACC well.ACC.PL bear.PRS.3SG
 “who treats you well (lit.: who treats you as a well-treated one)” (Yasht 13. 18)

22b) *yaθa* *hubərətō* *barāte* PASSIVE
 when well.NOM.SG.M bear.PRS.MDL.3SG
 “when he is treated well” (Yasht 10. 112)

[p. 41] In addition, some verbs are used only in the middle, among these “sit” and “wear” (23-24):⁴²

23) Avestan (Yasht 10. 45, Gershevitch 1959:96)

yeŋhe ... *vīspāhu* *pāti* *barəzāhu* ... *spasō* *āŋhāre*
 REL.GEN.SG.M all.LOC.PL to height.LOC.PL watcher.PL sit.PRS.MDL3PL
 “for whom on every height (...) watchers are sitting”

24) Avestan (Yasht 13. 3)

asmanəm ... *yim* *mazdā* *vaste* *vaŋhanəm*
 sky.NOM/ACC.SG REL.ACC.N PN.NOM wear.PRS.MDL.3SG garment.ACC.SG
 “the sky (...) which [Ahura] Mazda is wearing as [his] garment”

The more archaic among the Middle Ir. languages, Khotanese and Sogdian, preserve the Old Ir. voice distinction (25-26), again from Sims-Williams 1994.⁴³ Middles of the type “sit”, “wear”, etc., are also found (27). The Khotanese phrase “to remember” in (33) also always employs the middle of “do” (Emmerick 1987:287).

25) Khotanese

25a) *jsīndä* ACTIVE
 kill.PRS.3SG [+VOL, +INST, -AFF]
 “he kills [someone]”

25b) *uysāno* *jsande* REFLEXIVE
 self.ACC.SG kill.PRS.MDL.3SG [+VOL, +INST, +AFF]
 “he kills himself”

⁴² On these verbs, see Kellens (1984:19-23).

⁴³ For the middle in Khotanese, see also Canevascini 1991. For Sogdian cf. Yoshida (2009:299, 303).

- 26) Sogdian
- 26a) *wynt* ACTIVE
wēn-t [+VOL, +INST, -AFF]
 see.PRS-3SG
 “s/he sees”
- 26b) *wynty* INTRANSITIVE / PASSIVE
wēn-te [-VOL, -INST, +AFF]
 see.PRS-MDL.3SG
 “it appears / is seen”
- 27) Khotanese (Emmerick 2009:394f.)
- ysaiye* INTRANSITIVE / PASSIVE
 be born.PRS.MDL.3SG [-VOL, -INST, +AFF]
 "s/he is born"

However, only a few verbs are used both in the active and the middle. One of these is Khotanese *bar-*, but the voice opposition has been lexicalised to yield ACT “carry” vs. MDL “ride” (Emmerick 2009:391). In Sogdian, the middle is limited to the 3SG (Yoshida 2009:299). The morphological middle is out of function soon after the archaic stage of Middle Iranian.

3.4 The features and development of the middle voice

As already indicated above, one may say that the Old Ir. (and Proto-IE) voice opposition expresses whether or not the participant occupying the position of the grammatical subject is affected by the action or not. For instance, the subject of reciprocals and reflexives (19b, 20b) is [+AFF] and also either [+VOL] or [+INST]; the subject of the passive likewise shares the property [+AFF] and is also [-INST] (22b). While intransitives are “semantically diverse” and can have various semantic properties (Næss 2007:214), the subject of the intransitive in (21b), “the land”, could perhaps be described as falling within Næss’ (2007:89-93, 117) “volitional undergoer” ([+VOL, -INST, +AFF]) if it is viewed as a kind of person. [p. 42] Middle-only verbs typically “denote postures or states of body and mind” (Klaiman 1991:104) such as those in (23-24), surely permitting a classification of [+AFF] as well.⁴⁴

Indeed, the middle voice has often been described as implying affectedness cross-linguistically (cf. the observations by Næss 2007:22, 82, quoting Lyons 1968:363 and Kemmer 1993:130), thus also its name *ātmanepada* (“form referring to [acts for] one-self”) in classical Indian grammar. In addition of being [+AFF], the subject of a middle verb shows varying properties so far as volition and instigation are concerned. In some uses, it is clearly [+VOL] or [+INST] or both, sharing these properties with a prototypical agent as defined in Section 2.2. Its being [+AFF], on the other hand, connects it to the subject of a passive, although the

⁴⁴ This does not imply that all intransitives inflect in the middle (cf. Klaiman 1991:44).

latter (being a patient) is semantically opposed to an agent. This could quite well be connected to the observation (Næss 2007:90, referring to Kemmer 1993) that the middle is characterised by “a low degree of distinguishability of participants”, whereby properties usually attributed to either agent or patient may be “converging on the same participant” (ibid.). Subjects of middle verbs are thus not prototypical agents, and in Næss’ framework, it is only to be expected that one will find cross-linguistic variation as to which subjects (if any) with the property [+AFF] will be encoded as subject in a given language.

The specific range of the functions that the middle covers in ancient IE languages is not self-evident from a cross-linguistic point of view. While verbs such as “sit”, “wear”, “remember” and some verbs of movement commonly inflect only in the middle (languages such as Fula and Tamil patterning like ancient IE languages), reflexives and reciprocals are otherwise not necessarily included among the uses of the middle. For instance, reflexives and reciprocals require an additional affix in Fula (Klaiman 1991:47-82).⁴⁵ The passive featuring as a function of the middle is even less common in languages that have a middle voice.

The IE middle, then, covers uses for which in other languages different morphology would be required, where the middle is distinct from reflexive / reciprocal constructions and the passive. For such other languages, features in addition to [+AFF] are required to describe the subject of a middle verb. For the IE middle, however, [+AFF] appears to yield an economical description of the subject a middle verb.⁴⁶

Competing constructions arise as the middle voice loses ground in Middle Iranian. For instance, the middle is retained for some uses in Khotanese while novel ways of expressing the same contrast are arising alongside. In the potential construction, the distribution of *KAR* vs. *BAW* still encodes the opposition [+INST] vs. [+AFF]. The hypothesis suggested by these data is that pairs such as (1) presuppose the disappearance of the Old Iranian middle voice and the substitution of the voice opposition by the pair *BAW* and *KAR*.

4. CPs in Old, Middle and New Iranian

Having surveyed some of the framework potentially relevant for an assessment of the rise of the light verbs marking features such as [INST] and [AFF], we will turn to the study of the development of complex predicates in Persian itself. As will be seen in Section 5.1, opinions concerning the date of the rise of CPs in Persian have ranged from Old Iranian or even earlier to the times of Classical New Persian. Given this wide expanse, a short review of the evidence appears to be useful.

⁴⁵ For a study of the middle voice in Fula, see Kaufmann 2007.

⁴⁶ Note that the view that the IE middle has the function of marking a subject as [+AFF] also renders the question (and the ensuing difficulties, for which cf. Klaiman 1991:45f.) superfluous whether the middle voice represents a secondary function of the passive, an extension of the reflexive or something else.

4.1 Old Iranian

4.1.1 Data

The rather few articles that study Persian CPs diachronically (among these Sheintuch 1976, Fritz 2009 and Ciancaglini 2011, not citing each other) sometimes quote examples from Old and Middle Persian to support the claim that CPs, or constructions very much like New Persian CPs, were already used in earlier stages of the language. Indeed, from the viewpoint of New Persian, some Old Ir. sentences appear to look like CPs. This chiefly concerns the frequent Old Persian phrase in (28).

28) Old Persian (Dareios' inscription at Bisotun II. 34, Schmitt 1991:58, 2009:55)

<i>a-v-d-a</i>	<i>h-m-r-n-m</i>	<i>a-k^u-u-n-v</i>
<i>avadā</i>	<i>hamaranam</i>	<i>akunava</i>
there	battle.NOM/ACC.SG.N	do.IPF.3PL

“there they fought the battle”

[p. 43] However, *hamaranam KAR* is less CP-like than it might seem. In each occurrence of this phrase, the text mentions an uprising in some province against which Dareios sends an army; the clause in (28) then refers to the location and/or date where that specific battle took place. The meaning is thus rather “they did / carried out this (specific) battle at the place / date X” with a heavy verb *KAR* “carry out, perform”.

Other Old Persian examples that have been quoted in the literature are even more questionable, or ambiguous. Such instances are here exemplified by (29) (quoted by Ciancaglini 2011:15 to demonstrate the presence in Old Persian of “more or less fixed antecedents” of New Persian CPs). While the overall meaning is rather clear, it is not sure that *patipadam* is an “adverbial phrase” as assumed by Ciancaglini (2011:15, following Kent 1953); an alternative possibility would be “I made / rendered the lordship X (= feature of the lordship, *patipadam* a predicate)”, thence the interpretation of the sentence remains uncertain.⁴⁷

29) Old Persian (Dareios at Bisotun I. 61f., Schmitt 1991:29, 53, 2009:45)

<i>x-š-č-m</i>	...	<i>a-v</i>	<i>a-d-m</i>	<i>p-t-i-p-d-m</i>	<i>a-k^u-u-n-v-m</i>
<i>xšaçam</i>	...	<i>ava</i>	<i>adam</i>	<i>patipadam</i>	<i>akunavam</i>
lordship.NOM/ACC.N		DEM.NOM/ACC.N	I.NOM	?	do.IPF.1SG

“the lordship (which...), that one I reestablished
(i.e.: I reestablished the lordship which...)”

In Ciancaglini's second Old Persian example (ibid.), she translates (following Kent 1953:142) *gāθavā akunavam* (30) as “I put in its place”, but this is at variance with the Babylonian version of the inscription according to Schmitt (2009:126), who follows Hoffmann

⁴⁷ Cf. Schmitt (2009:45).

([1955] 1975:55) in interpreting *gāθavā KAR* as “do in the right way”, thus again with a heavy verb.

30) Old Persian (Dareios’ inscription at Susa e 43f., Schmitt 2009:126)

<i>t-y</i>		[<i>p-r^u-u-v-m</i>	<i>n-i-y</i>]	<i>g-a-θ-v-a</i>	<i>k-r-t-m</i>
<i>taya</i>		<i>paruvam</i>	<i>naiy</i>	<i>gāθavā</i>	<i>kərtam</i>
REL.NOM/ACC.N	earlier	not	right	do.PP.NOM/ACC.N	
[<i>a-v</i>		<i>a-d-m</i>		<i>g-a-θ]-v-a</i>	<i>a-k^u-u-n-v-m</i>
<i>ava</i>		<i>adam</i>		<i>gāθavā</i>	<i>akunavam</i>
DEM.NOM/ACC.N	I.NOM	right		do.IPF.1SG	

“the one which earlier was not done properly, that I did properly
(I repaired buildings which were not built properly)”

For Avestan, Thordarson (2009:80f.), following Sheintuch 1976, quotes sentences with two accusatives such as (31), assuming that these could be predecessors of New Ir. CPs.

31) Avestan (Videvdad 4. 40 etc., Thordarson 2009:80, quoting Sheintuch 1976)

<i>yō</i>	<i>narəm</i>	<i>frazābaodaŋhəm</i>	<i>snaθəm</i>	<i>ǰānti</i>
REL.NOM.SG.M	man.ACC.SG	mortal.ACC.SG	blow.ACC.SG	strike.PRS.3SG

“who strikes a man a mortal blow”

However, an interpretation as a CP is by no means compelling, since “strikes, hits” certainly has its full value here. The same applies to the other Avestan examples adduced by Sheintuch and Thordarson. Also, double accusatives are quite common in ancient Indo-European languages, and are a phenomenon quite separate from CPs.⁴⁸

4.1.2 Evidence from Aramaic

Aramaic data perhaps supplement the meagre evidence for possible CP-like structures in Old Iranian. According to Ciancaglini (2011:15f.), some Aramaic documents dating to the 5th century BC show nouns combined with the Semitic root ‘*BD* “do” in phrases such as (32). Ciancaglini (2011:9) notes that the occurrence of such patterns in Semitic is particularly noteworthy since these languages usually employ the so-called *pa^{cc}el* forms for denominatives, and combinations such as (32) are foreign to the structure of Semitic.

[p. 44]

32) Aramaic (Ciancaglini 2011:15)

<i>hndrz</i>	<i>y^cbdwn</i>
instruction	do.IPF.3PL.M

“they will give instructions”

⁴⁸ Cf. Reichelt (1909:228f.) and Skjærvø (2009:106). For parallels in Latin and Greek, see Telegdi (1951:324).

Judging from the instances cited by Ciancaglini, at least some of the nouns featuring in the examples are Iranian words (such as *handarz* in (32)), confirming the assumption of the pattern being copied from Iranian. Since the relevant documents originate from Susa, the language in question is likely to be Old Persian or an Ir. dialect close to it.

As argued by Ciancaglini, instances such as Aramaic *hndrz y^cbdwn* in (32) appear to imply the existence of corresponding Old Persian phrases (in this case, the expression **handarzam* (?) *KAR*, not attested in Old Persian itself) on which the Aramaic examples are modelled. If this is so, such patterns would have been more common in (some variety of) Old Persian than it appears from the texts themselves. Following a thought suggested for Middle Persian by Fritz (2009:35, 50), it seems possible that spoken Old Persian made more use of such patterns than the style in which the inscriptions were composed, the latter certainly being the most elevated register available to the authors.

4.1.3 Summary

The evidence for CPs in Old Iranian seems rather meagre. Old Persian examples with *KAR*, insofar as they are clear at all, appear to point to *KAR* being used as a heavy verb without semantic bleaching or other characteristics commonly attributed to light verbs. The argument structure of *KAR* does not seem to be reduced either, nor is this the case for verbs in double accusative constructions such as the one employed in (31).

However, Aramaic might indicate that (some variety or register of) Old Persian did show patterns composed of a noun and *KAR* that may have been fixed expressions of some type. An example such as (32), or rather its presumed Old Persian equivalent (with *KAR*) would not necessitate an interpretation as CP yet, but it could be seen as being on the first steps on its way to become one, showing a somewhat generalised meaning translatable as “carry out, perform”. In the light of the later stages, this may be seen as the start of a development leading to a light verb, which is defined by (inter alia, see Section 2.2) as having “semantics (...) which are not very specific” by Bown (2008:163). Still, even as “perform”, *KAR* retains the semantic features [+INST], [+VOL] and possibly [-AFF], thus prototypically transitive by the criteria of Næss 2007.

4.2 Middle Iranian

4.2.1 Data

The issue of CPs in the Middle Iranian period has hardly been investigated, but examples such as (33) can be found already in archaic Middle Ir. languages. Khotanese *byāta* originally is a perfect participle (“remembered”), but has become a fixed and uninflected form combined with *KAR* (“remember”) and *BAW* (“be remembered”).⁴⁹

⁴⁹ See Emmerick (1987:278). Technically, *byāta* would be the NOM/ACC.PL (M/N) or NOM.SG.F.

33) Khotanese (Emmerick 1987:278)

33a) *balysi ... yan-de ṣṣahāne ... byāta*
Buddha.GEN/DAT do.PRS-MDL.3SG virtue.ACC.PL remembered
“he remembers the Buddha’s virtues”

33b) *ko ju aysu byāta hām-āne*
SUB PTC I remembered become.PRS-MDL.SBJ.1SG
“may I be remembered!”

So far as Middle Persian is concerned, the situation is surprisingly unclear. On the one hand, grammars of Middle Persian have stated the presence of CPs in Middle Persian such as *zēn kardan* “to saddle” and *rōšnīh kardan* “to illuminate” (lit. “saddle / light do”).⁵⁰ However, as noted by Fritz (2009:8), the tendency is that the same few examples keep being quoted, and that, if referenced at all, they come from the same small selection of texts such as the *Kārnāmag ī Ardašīr*, i.e. from late texts, thus liable to show New Persian influence. Such evidence cannot be regarded as representative nor compelling.

[p. 45] Some examples quoted by Brunner (1977:22) and Ciancaglini (2011:13f.)⁵¹ from the older layers of Middle Persian look more promising, of which (34-35) appear to be the most likely candidates for CP-hood. However, they still seem to permit an alternative interpretation as noun plus heavy verb such as “I will make an advice (i.e. provide a lesson) for them” in (34).

34) Middle Persian (Psalter, Psalm 131.12)⁵²
ZNH gwk²dyhy ZYšn ²ndlcy ^cBYDWNm
ēn gugāyih i=šān handarz kun-am
DEM testimony DEM=PRO.3PL advice do.PRS-1SG
“these testimonies which I will teach them”

35) Middle Persian (Shapur II’s inscription at Persepolis, line 10, 4th c. A.D.)⁵²
Ṗš šhpwḥry MLK²n MLK² Ṗlyny krty
u=š šāpuhr šāhān šāh āfrīn kard
and=PRO.3SG PN king of kings blessing do.PST
“he blessed Šāpuhr, the king of kings”

In (35), the context reads (Frye 1966:85): “He caused great rejoicing, and ordered rites performed for the Gods. He gave blessings to his father and ancestors. Then he offered blessings to Shapur the King of Kings, to his own soul, and also to him who built this structure [= building]”, in all instances with the same *āfrīn kard* “give / offer blessings”. Rather than an interpretation along the lines of a New Persian CP, it appears more plausible to me that the

⁵⁰ Thus e.g. Rastorgueva / Molčanova (1981:135f.), Sundermann (1989:152), Skjærvø (2009a:214).

⁵¹ Most of Ciancaglini’s Middle Ir. examples are those also quoted by Brunner 1977.

⁵² Brunner’s (1977:23) transliteration and translation with Ciancaglini’s (2011:13) transcription.

text refers to a ritual practice involving the recitation of a prayer for the persons quoted (thus *āfrīn KAR* not a general “to bless”, but a technical or even liturgical term). Taking together all instances of *āfrīn KAR* in the Middle Persian inscriptions,⁵³ a translation as “perform a blessing ritual / prayer” seems to fit best, suggesting an interpretation of *KAR* as heavy verb with *āfrīn* as its direct object. On the other hand, as discussed for the Aramaic evidence above (4.1.2), the meaning “perform” does appear to be a somewhat wider sense than “make, do”, and would also seem quite adequate for (34). The dative argument =*šān* in (34) does not presuppose a particular CP argument structure either, as it is liable to an interpretation “I performed a testimony for them”, thus a common benefactive.

4.2.2 Evidence from Syriac

Following up on the Semitic evidence for CPs in Iranian, Ciancaglini (2011:14f., 2006) says that Syriac shows constructions combining ‘*BD* “do” with a noun – either an Iranian noun (e.g. *nhšyr*’ *BD* “hunt (lit. do hunt)”), a Greek one, or an inherited Semitic word as in (36).

36) Syriac (Ciancaglini 2011:14f.)

ḥwʳ’ *BD* / *ḥawwar*
 white do make.white
 “to whiten”

She argues that this pattern, which is “attested from at least the fourth century A.D. onwards”, “is most likely a calque on Middle Persian”. Syriac data would thus confirm the use of *KAR* in Middle Persian in phrasal patterns of the type “provide an advice, perform a prayer”, etc.

4.2.3 Denominatives and causatives

One important point for the evaluation of CPs in Middle Persian is the status of denominatives and causatives. In fact, all the verbs which Megerdooian (2012:183) quotes to illustrate that “the majority of simple verbs in Middle or Classical Modern Persian have now been replaced by light verb constructions” are transparent denominatives, not the products of Old Ir. verb stems.⁵⁴ The replacement referred to thus merits a closer look.

As noted by Ciancaglini (2011:11f.), following Telegdi (1951:320-322), Middle Persian phrasal patterns coexist with causatives / denominatives, both with and without suffix *-ēn-* (less frequently *-ān-*), yielding series such as (2’).

[p. 46]

⁵³ For a list of attestations, see Gignoux (1972:16).

⁵⁴ Megerdooian (ibid.) also notes an ongoing process of substitution of CPs for the remaining simple verbs (which are not denominatives).

- 2') Middle Persian⁵⁵
- 2a) *nām kardan* / *nām-ēn-īdan* / *nāmīdan*
 name do.INF name-CAUS/DENOM-INF name-INF
 “to name, to call”
- 2b) *handarz kardan* / *handarz-ēn-īdan*
 advice do.INF advice-CAUS/DENOM-INF
 “to give advice”

Sundermann (1989:152, likewise following Telegdi) even describes the CP-like pattern with *KAR* as periphrastic expressions of denominatives with causative value. So the replacement under discussion is the ousting of one type of denominative, the morphological one, by another one, viz. the CP pattern (Telegdi 1951:320f., Pirejko 1975:336, Ciancaglini 2011:11f.), implying a shift in productivity between two competing ways of forming denominatives. As a result, formations with *-ēn-* have disappeared from Modern Persian, while suffixless denominatives remain productive at a low level, permitting formations such as *raqṣīdan* “to dance” from (Arabic) *raqṣ* “dance”.⁵⁶

The suffix *-ēn-* is likely to contain the Proto-IE suffix **-eǵe-*, which forms both causatives and denominatives.⁵⁷ It thus appears that the rise of CPs is linked to the history of morphological causatives, which happened to be denominatives at the same time.

4.2.4 Summary

While the status of CPs in Middle Persian (and Middle Iranian in general) is not clear, CPs appear to gain some prominence in Middle Persian. Specifically, it seems that the denominative / causative suffix *-ēn-* played a role, so that verbs with this suffix and CPs are competing constructions. I argue that this situation is a crucial point in the development of Persian CPs: two alternative verbal patterns are available to the speakers, both with causative value, one the morphological one with the suffix *-ēn-*, the other one with *KAR*. It seems plausible, then, that the pattern that later on became the CP construction entered the system at the point where the meaning is a causative one. If this proposal is on the right track, it suggests that the rise of CPs could not have happened before the Middle Persian period, as the suffixes *-ēn-* and *-ān-* are Middle Ir. innovations.⁵⁸

Middle Persian examples for CPs quoted by grammars often come from texts that are liable to show New Persian influence. So far as examples from earlier Middle Persian texts are concerned, it seems that *KAR* does not yet show particular semantic bleaching, but perhaps

⁵⁵ (2a) with *nāmīdan* from Sundermann (1989:152), the others from Ciancaglini (2011:11f.).

⁵⁶ Thus Telegdi (1951:320f.), Sheintuch (1976:145 n. 3), contra Ciancaglini (2011:3).

⁵⁷ Ciancaglini (2011:11f. n. 29), Henning (1934:212-214). It is possible that *-ān-* is originally the Parthian causative suffix (Henning 1934:213) while *-ēn-* appears to be specifically Middle Persian.

⁵⁸ Sundermann (1989:151), referring to Salemann (1901:305), who says: “The formation consisting in adding the causative suffix *-ēn-*, much more rarely *-ān-*, to the present stem only emerges in Middle Persian and is undoubtedly of denominative origin.”

some generalised semantics interpretable as “perform” (and the Syriac example (36) is in fact rather close to the meaning of the heavy verb “make”). *KAR* thus seems to retain the semantic features [+INST], [+VOL] and possibly [-AFF], agreeing with the features of the suffix *-ēn-* with which it is in competition.

So far as argument structure is concerned, studies are likewise lacking, but Jügel (2012/I:293) suggests that the degree of lexicalisation does not affect argument structure, i.e. any Middle Persian combination of noun and transitive verb would have its agent in the oblique case (insofar as case marking is preserved in Middle Persian) even if the overall semantics is intransitive (i.e. patterns such as New Persian *zamīn x^wordan* “to fall (lit.: eat/hit earth)”).

4.3 New Iranian

For the stage of New Iranian, it seems suitable to discuss Persian in the context of other Ir. languages, and in the light of neighbouring languages, to get a picture of possible developments.

4.3.1 Light verbs in various New Ir. languages

Throughout Iranian, *KAR* “do” for [+INST] and *BAW* “be / become” for [+AFF] meanings are by far the most commonly employed light verbs (labelled “dummies” by Windfuhr 1979:113). A few examples are given in (37).⁵⁹

[p. 47]

37) CP pairs in New Ir. languages (selection)

	[+ INST]			[+ AFF]		
Kurmanji	<i>ava</i> thriving	<i>kirin</i> do.INF	“to build”	<i>ava</i> thriving	<i>bûn</i> become.INF	“to prosper”
Gilaki	<i>ĵam</i> collecting	<i>kudān</i> do.INF	“to collect”	<i>ĵam</i> collecting	<i>bostān</i> become.INF	“to assemble (ITR)”
Yaghnobi	<i>ĵudō</i> separated	<i>kun-</i> do	“separate”	<i>ĵudō</i> separated	<i>vī-</i> become	“become separated”

In Ossetic complex predicates, the distinction is visible only in the perfective aspect while the imperfective aspect uses *KAR* for both transitive and intransitive meanings (38).⁶⁰

⁵⁹ The *caveat* mentioned in note 3 applies here as well. For CPs in Kurmanji and Sorani Kurdish, cf. Bedir Khan / Lescot (1970:183-185, thence the examples) and Blau (1980:113-117), respectively. The Russian grammars of Kurdish also quote a large selection of CPs, see e.g. Kurdoev (1978:138-143). For CPs in Gilaki, see Rastorgueva et al. (1971:132-135, examples from p. 133), and cf. Bird (2007:84f.) on CPs in Yaghnobi. For a list of Balochi CPs with “do” and “eat”, see Farrell (1995:232f., 2008:126f.), respectively.

⁶⁰ Cf. Fritz (2009:37f.). Aspect is chiefly marked by prefixes in Ossetic, which are used for the perfective aspect while the verb without prefix encodes the imperfective aspect in the past and future tenses; the present tense is neutral (cf. Abaev 1964:45). For the transitive past, see Section 5.3.

38) Ossetic (Abaev 1964:67f.)

	PRS		PST.TR		PST.ITR
38a)	<i>mæstʃ</i>	<i>kænin</i>	<i>s=mæstʃ</i>	<i>kod-ton</i>	<i>s=mæstʃ</i> <i>dæn</i>
	angry	do.INF	PFV=angry	do.PST-TR.1SG	PFV=angry COP.1SG
	“to anger, to be angry”		“I angered”		“I got angry”
38b)	<i>yu</i>	<i>kænin</i>	<i>ba=yu</i>	<i>kod-ton</i>	<i>ba=yu</i> <i>dæn</i>
	unite	do.INF	PFV=unite	do.PST-TR.1SG	PFV= unite COP.1SG
	“to unite, to be united”		“I united”		“I was united”

While “do” and “become” are also common light verbs in other IE languages (cf. e.g. Fritz 2009:45f. for Sinhalese) and beyond,⁶¹ *KAR* and *BAW* are not the only ones used in Iranian. For instance, Ossetic also employs, besides *kænin* “to do”, also *lasin* “to draw, carry”, *darin* “to hold”, *kæsin* “to look”, *marin* “to kill” and *mælin* “to die” (cf. Abaev 1964:67, 69 and Thordarson 2009:77). (39) shows data from further contemporary Ir. languages.⁶² The evidence of languages such as Ossetic and Zazaki, which have not been under heavy Persian influence in the course of their history, is particularly relevant.

39) Light verbs in selected New Ir. languages

	[+INST]		[+AFF]	
Pashto	<i>kawəl</i>	“to do”	<i>kədəl</i> (IPFV) / <i>šwəl</i> (PFV)	“to become”
	<i>larəl</i>	“to hold”	<i>axistəl</i>	“to take”
			<i>kšəl</i>	“to pull”
			<i>mindəl</i>	“to find”
			<i>xwaṛəl</i>	“to eat”
Wakhi	<i>tsar-</i>	“do”	<i>wots-</i> (past <i>vit-</i>)	“become”
	<i>ḍir-</i>	“hold”		
	<i>goḥ-</i>	“make”		
	<i>xaš-</i>	“pull”		
	<i>di-</i>	“beat, hit”		
Zazaki	<i>kerdiş</i>	“to do”	<i>amiyayiş</i>	“to come”
	<i>dayiş</i>	“to give”	<i>şiyayiş</i>	“to go”

[p. 48] The data suggest that, in addition to *KAR*, the verbs “hold” (cognates of New Persian *dāştan*) and “give” (New Persian *dādan*) are in wide use for the [+INST] pattern

⁶¹ For parallels in Turkic and Caucasian languages, see Thordarson (2009:79).

⁶² For Pashto CPs, see the lists in Lorenz (1979:86-89); cf. also Fritz (2009:36); for Wakhi, see Bashir (2009:833). For detailed discussion of CPs in Zazaki, see Paul (1998:100f., 131-133, 2009:556); “become” appears to be chiefly used in CPs whose nominal element is a Turkish participle in *-miş*, e.g. *dışmiş biyayiş* “to think (lit. thinking become)” (ibid.).

while various verbs of movement are added to *BAW* for the [+AFF] one.

So far as the [+INST] pattern is concerned, this picture appears to agree rather well with data from earlier stages of New Persian. As pointed out by Fritz (2009:16ff.), *KAR* remains the single frequent light verb in Early New Persian while other light verbs became common only later on, cf. the pairs in (40).⁶³

40) Persian (Fritz 2009:17)

40a) Early New Persian (*Shahnama*)

<i>pōziš</i>	<i>kardan</i>	<i>pāsux</i>	<i>kardan</i>
apology	do.INF	answer	do.INF

40b) Modern Persian

<i>pūzeš</i>	<i>x^vāstan</i>	<i>javāb</i>	<i>dādan</i>
apology	want.INF	answer	give.INF
“to apologise”		“to answer”	

4.3.2 Evidence from Armenian

At this point, the evidence from non-Iranian languages that have been under heavy influence from Persian (and closely related languages), comes in again, since Armenian appears to provide evidence for a stage with a particular inventory of light verbs.

Discussing Armenian phraseology and idioms copied from Iranian, Schmitt (1983:103f.) points out that complex predicates occupy a special place: “particularly numerous are expressions with the auxiliaries *aṛnel* “make”, *linel* “be”, *harkanel* “beat”, *ownel* “have [also: hold, possess]”, and *tal* “give”; there is even a series of such combinations with *aṛnel* or *harkanel* which, as shown particularly by New Persian material, correspond to expressions with *kardan* “make” or *zadan* “beat.” Schmitt continues to say that there are Armenian denominatives of the same meaning beside many of these. Many of the examples cited by Schmitt involve an Iranian noun, but others have an Armenian one, as shown in (41) by *dat* and *owhił*, respectively, in a CP corresponding to Persian *dād kardan*.

41) Armenian (Schmitt 1983:104)

<i>dat</i>	<i>aṛnel</i>	/	<i>owhił</i>	<i>aṛnel</i>
judgment	do.INF		judgment	do.INF
“to judge (deliver a judgment)”				

Evidently, the question is from which Iranian language, and at which time point, Armenian imported the system. One obvious option would be the early stage of Parthian, which is the chief source of lexemes borrowed from Middle Iranian (i.e. the vast majority of the items preserving postvocalic voiceless stops, as is the case for *dāt*). The number of Parthian words in the Armenian lexicon presupposes a wide-spread bilingualism for the

⁶³ Telegdi (1951:321 n. 24).

Arsacid period (for a survey see Schmitt 1983:73f.), obviously a condition for the adoption of a structural feature such as complex predicates. However, it is by no means sure that Parthian is the source, and the presence of Parthian nouns in Armenian CPs could also be due to the fact that the former are an integral part of the Armenian lexicon and once they are adopted can of course be used just like inherited nouns in any structure once they have been adopted. Moreover, the Ir. parallels cited by Schmitt are rather from Persian than from Parthian. This may also be due to the fact that complex predicates in Parthian are even less studied than those in Middle Persian, but, as in fact pointed out by Schmitt, the actual evidence for Iranian parallels comes from New Persian.

At any rate, the specific group of verbs functioning as light verbs in Armenian complex predicates is noteworthy. It appears to point to the existence of a stage with “be” as the only light verb to encode [+AFF] while there are several verbs besides “do” for the [+INST] slot (“hit”, “give”, “hold/have”), presupposing a (Persian?) system with *BAW* on the one hand and *KAR*, *zadan*, *dādan* and *dāštan* on the other. Armenian would thus provide evidence for the moving into the system of several items marking [+INST] in complex predicates. This might at the same time be further evidence for the primary role of the [+INST] pattern in CPs, and for *KAR* being established as the first light verb.⁶⁴

[p. 49]

4.3.3 Verbs of movement

Turning now to the verbs seen in the [+AFF] slot in (39) in addition to *BAW*, verbs of movement appear to play a particular role. Again, the evidence of Zazaki is particularly precious because the language has not been under strong Persian influence since it became located in Eastern Anatolia. Zazaki uses “come” and “go” for [+AFF] CPs, etymologically cognate to Persian *āmadan* and *šodan*. In Persian, the latter has undergone a semantic shift from “move forward” to “become”. An intermediary stage of this development may be seen in Early New Persian examples such as (42).

42) Early New Persian (*Shahnama*; Fritz 2009:17)

hamē jāng u mardī firō šud

entirely battle and manliness down go.PST

“battle and manliness have all been lost (gone astray)”

The same group of [+AFF] verbs are used as auxiliaries for the analytical passive voice that arises after the Old Iranian stage. Sogdian, Parthian, Middle Persian and Balochi, among others, use *BAW* while other Ir. languages employ verbs of movement: “come” (Old Ir. **ā-i- / ā-g(m)ata-*, Persian *āmadan*) forms the passive in Kurmanji, “move forward, go” (**čyaw-*, Persian *šodan*) is used in Pashto, some Pamir languages and New Persian; and Khotanese and

⁶⁴ Ciancaglini (2011:9f.) also notes the presence of CPs in modern Semitic languages such as Neo-Aramaic and Anatolian Arabic, where CPs are motivated by Persian and Kurdish in the first case and by Kurdish and Turkish in the latter. However, this could be a rather recent phenomenon.

Ossetic use both “become” and “move forward”.⁶⁵

Verbs of movement moving into the system of CPs recalls the fact that these verbs are often treated as middle verbs in Indo-European and cross-linguistically (see Sections 3.3, 3.4), and employ *BAW* in the potential construction in those Iranian languages that show this pattern (cf. Section 3.1).

5. Evaluation and conclusion

5.1 Owing to the scarcity of available studies, opinions about the status and development of CPs in stages prior to New Persian have diverged considerably. For instance, Thordarson (2009:80f.) concludes from examples such as (31) that CPs “existed *in nuce* already in early Common Iranian” and are therefore “a native inheritance, rooted in prehistoric Iranian”. Going even further, Ciancaglini (2011:4) considers CPs “an Indo-Iranian feature”. Fritz (2009:8f., 35), while dismissing such an early date, assumes that large-scale areal influences (Near Eastern – Caucasian) caused a “predisposition” of Indo-Iranian languages for CPs, which was realised later on, when language changes had made their morphosyntax receptive to respond to that predisposition.

Conversely, other authors have assumed that CPs are a rather late phenomenon.⁶⁶ Sheintuch (1976:140) concludes from the limited number of Middle Persian CPs cited in modern descriptions (sic; not: found in the texts) that CPs were either not productive, or not uniform enough to be structurally relevant. Still others have assumed that the rise of CPs is a result of the advent of Arabic – an opinion going back at least as far as Kazimirski (1883:66), who claims that “in order to remedy this impoverishment of their language [by the supposed disappearance of simple verbs], the Persians have deemed it necessary to form verbs with the help of Arabic action nouns joined to Persian verbs [such as *kardan*]”. However, this is at variance with the data. As pointed out by Telegdi (1951:330f.), the processes of the rise of CPs and the substitution of Persian words by Arabic ones are independent of each other. In fact, “Arabic nouns are not borrowed to be used for the formation of periphrastic verbs [= CPs], but they are used [for CPs] because they were borrowed” (Telegdi 1951:330f.). Indeed, a CP using an Arabic noun “is based on an older one employing an Iranian one” in many cases (Telegdi *ibid.*, cf. also the data from the *Shahnama* etc. quoted by Fritz 2009:17, 50). This said, “the high frequency of CPs in Persian [today] is surely to be explained by the “Islamisation” hypothesis, but the existence of the category [= the CPs] is not” (Fritz 2009:35, similarly Sheintuch 1976:140). Instead, as also suggested by the argument in 4.2, the replacement of simple verbs by CPs is thus likely to be a gradual process dating back to Middle Persian (cf. Windfuhr 1979:114).

5.2 I argue that the evidence discussed in Section 4. suggests the following line of development: Old Persian appears to have shown phrases containing a noun and *KAR*. While

⁶⁵ See Emmerick (2009:397) and Axvlediani (1963:264), respectively, for these languages. In Zazaki, the passive is synthetic (cf. Paul 1998:95f., 2009:553f.).

⁶⁶ Folli et al. (2005:1369) even date the substitution of CPs for simple verbs to as late as the 13th century.

the status of this pattern appears marginal at best in the Old Persian texts, it seems to have been clear enough in other styles to motivate some loan translations in certain Aramaic documents discussed by Ciancaglini 2011. In view of the later development, it seems plausible, then, that the existence of combinations of a noun and *KAR* such as in (28) may have played a role for the later development even if at this stage *KAR* is not a light verb by criteria such as reduced argument structure or deficient semantics. It does seem, though, that it is beginning its way towards a generalised meaning, losing the aspect “produce” that is one component of the meaning “make”, and moving into the direction of “perform, carry out”.

[p. 50] Middle Persian shows some more examples looking like proto-CPs, but, owing to the absence of specialised studies, the status of these within the language is unclear. At any rate, the pattern of noun plus *KAR* was apparently sufficiently prominent to be copied into Syriac. There was also an association of this Middle Persian pattern with morphological denominatives. These were formed with the suffix *-ēn-* that at the same time had a causative meaning. This suffix is a Middle Persian innovation, which sets a *terminus post quem* for the subsequent replacement of these morphological causative-denominatives by light verb constructions of the same function. It is possible that at this stage (within the Middle Persian period) CPs with *KAR* might have acquired a status similar to structures found in German, Hungarian, etc., the CP competing with a full verb of the same meaning similar to German *Antwort geben* with *antworten* (cf. Section 2.1).

The association taking place in Middle Persian between CPs with *KAR* and causatives / denominatives suggests that at this stage, *KAR* had developed a still more general meaning, which broadly seems to equal the verbal expression of the features [+INST], [+VOL] and [-AFF] that per Næss 2007 (see Section 2.2) are the characteristics of a prototypical agent of a transitive verb.

Agreeing with the causative argument just made, evidence from Aramaic and Syriac appears to suggest that only the construction with *KAR*, and not (also) a corresponding one with *BAW*, was prominent enough in Old and Middle Persian to be copied into neighbouring languages. This appears to be linked to the history of the middle voice in Iranian: only with its disappearance in early Middle Iranian the scene appears to be set for the entrance of a periphrastic expression mirroring *KAR* and covering the range of the IE mediopassive, encompassing typically middle verbs such as “sit”, “wear” and verbs of movement.

At this point, a look at the so-called potential construction seems doubly relevant: the pattern using *KAR* and originally meaning (according to Benveniste) “succeed, achieve” is a further point in establishing this verb as indexing causative and [+INST] meanings. Secondly, the potential auxiliary *KAR* is used in a pair with *BAW* in a distribution paralleling the IE voice opposition of active vs. mediopassive.

Comparative data from CPs in various contemporary Ir. languages favour the assumption that the verb pair *KAR* and *BAW* lies at the heart of CPs as a structure. As pointed out by Bovern (2008:171), “once the complex predicate structure exists in the language, other items may be recruited into the construction.” The first decisive step appears to have been the opening of a periphrastic expression of causation to include a parallel

pattern covering the range of the Old Iranian middle voice as this disappeared. So the CP system appears to have had the slots of [+INST] (encoded by *KAR*) and [+AFF] (with *BAW*). This being the core of the system seems to be confirmed by the fact that those Ir. languages having etymologically unrelated verbs for “do” or “become” still show the same opposition.

The evidence of Armenian and of contemporary Iranian languages appear to suggest that next was a stage integrating more light verbs into the [+INST] slot, among these the equivalents of New Persian *dāštan* “have (originally: hold)” and *dādan* “give (originally also: put)”. With view of the combined evidence of New Iranian languages, verbs of movement are likely to have joined *BAW* for the [+AFF] pattern then.

5.3 The verbs just mentioned are all also used as auxiliaries. This was already shown for the verbs of the [+AFF] slot (Section 5.1), but it is also true for some of the [+INST] slot, viz. the equivalents of Persian *dāštan* and possibly *dādan*.

In several Eastern Ir. languages, “hold / have” (Old Ir. *dār-*) is used to form a periphrastic perfect of the type parallel to the Germanic and Romance “have” vs. “be” perfect, here exemplified by Sogdian (43).⁶⁷ As pointed out by Wendtland 2011, the form with “hold / have” gradually becomes more frequent within the history of Sogdian: beginning in the stage reflected by the Manichean texts and more systematically in the Christian texts (where the final *-t* of the past stem is often lost in contact with the *δ-* of the auxiliary), “hold, have” is also used for “unergatives” while “unaccusatives” continue to take the copula.⁶⁸

43) Sogdian (Yoshida 2009:301)

TR	<i>wyt</i>	<i>δʳ²m</i>	ITR	<i>ʷyt²ym</i>	
	<i>wēt</i>	<i>δār-ām</i>		<i>āyat</i>	<i>=im</i>
	see.PP	hold/have.PRS-1SG		come.PP	=COP.1SG
	“I saw”			“I came”	

[p. 51] As in the potential construction, Xwarezmian generalises the transitive pattern for all verbs (44).⁶⁹

⁶⁷ In Khotanese, the ITR perfect is formed in the same way as in Sogdian and Ossetic, but the TR perfect uses the copula with an enlarged form of the perfect participle, for which various derivations have been suggested (e.g. a participle in **-tāvant-*, Sims-Williams 1997:322-324, differently Christol 1990:43 and others).

⁶⁸ Cf. Sims-Williams (1989:189), Yoshida (2009:301). Terminology for the Sogdian pattern has varied. Using “preterite” for combinations involving the past stem (following Yoshida 2009:302f.) has the advantage that “perfect” is then free for combinations of the perfect participle (= past stem + **-aka-*, which agrees with the TR object and the ITR subject). The ergative construction of the TR past / perfect is also found in Sogdian, and seems to be an older usage than the preterite (Yoshida 2009:302, 325). Note that Yaghnobi does not show such a construction, but uses the past participle plus enclitic pronouns for the TR perfect. The parallelism of the Sogdian construction and the perfect forms of Romance and Germanic has been pointed out by various authors since Benveniste (1929:48f.). Likewise parallel is the Hittite perfect, using another etymologically unrelated verb “hold / have” vs. “be” in a very similar distribution (cf. Wendtland 2011:41f.).

⁶⁹ Durkin-Meisterernst (2009:353f.).

44) Xwarezmian (Durkin-Meisterernst 2009:354)

44a) *ktk* *δ²ry²y*
aktēk *δāray-āni*(?)
 do.PFP hold.PRS-SBJ.1SG
 “I should have made”

44b) *ȳdk^o* *δ²ry-d.* *^omy*
ēdek *δāray-eda-* ?
 go.PFP hold.PRS-IPF.3SG PTC(?)
 “she had gone”

Ossetic likewise has a double construction for the preterite (45, 38) which has been described as showing that “the category of transitivity is expressed morphologically” (Thordarson 1989:474). However, intransitive meanings being grouped with passives, and the formation being called “middle” (“Imperfectum medii”) by Miller (1903:72) suggest that the pattern, combining the past stem with the copula, in fact rather represents [+AFF].

45) Ossetic

	PRS	PST.TR	PST.ITR
45a)	<i>xiz-in</i>	<i>xis-ton</i>	<i>xistæn</i>
	graze.PRS-INF	graze.PST-TR.1SG	graze.PST.COP.3PL
	“to graze (TR/ITR)”	“I grazed (TR)”	“I grazed (ITR)”
			(Thordarson 1989:474)
45b)	<i>kal-in</i>	<i>kald-ton</i>	<i>kaldst#</i>
	pour.PRS-INF	pour.PST-TR.1SG	pour.PST.COP.3PL
	“to pour (TR/ITR)”	“I poured”	“they are poured”
			(Abaev 1964:53f.)

For the transitive pattern, a derivation of the Ossetic transitive preterite from “hold” would be attractive in the light of the Sogdian and Xwarezmian formations, but it would imply an irregular loss of the *r* vs. the preserved *dar-* “hold”. While it seems difficult to exclude the possibility of such a dual development, an alternative possibility is the derivation from the Proto-IE root **d^heh₁* (Ir. **dā*) “put, place” (Christol 1990:44, Lubotsky apud Cheung 2002:140). For the semantics, Christol compares Latin *facio* and Slavic *děti* from the same root, to which one could add English *do*, German *tun*, etc. Also morphologically, the formation would be parallel to the weak preterite of Germanic.⁷⁰

Table (46) summarises the results.

⁷⁰ Christol (1990:44) further assumes that the “analogy between Gothic and Ossetic is easily explained by a long period of symbiosis”, which appears rather speculative. For the Germanic weak preterite, see e.g. Krahe / Meid (1969:127-129).

46) Auxiliary verbs in Iranian (selection)

	[+INST] patterns			[+AFF] patterns	
<i>KAR</i> “do”	transitive potential	(many Ir. languages)	<i>BAW</i> “be, become”	middle and passive potential	Khotanese, Sogdian, Balochi, Pashto
<i>dār</i> “hold, have”	transitive past	Sogdian, Xwarezmian		intransitive past	Sogdian, Ossetic
<i>dā</i> “put; give”		Ossetic		compound tense forms (perfect, past perfect)	(general)
				passive	(many Ir. languages)
			<i>*čyu</i> “move forward”	passive	Khotanese, Pashto, Ossetic, New Persian, etc.
			<i>*ā-gam</i> “come”		Kurmanji

[p. 52]

5.4 The link apparently shown by Iranian light verbs and auxiliaries merits further discussion. As argued by Butt (2010:49-51), an auxiliary, defined as a mere “functional item along the lines of tense and aspect auxiliaries”, is to be distinguished from a light verb since both the light verb and the non-verbal element contribute something to the predication of the CP. On Butt’s approach (2010:53f., 64f.), auxiliaries differ from light verbs insofar as they are instances of a grammaticalisation process, often developing to bound morphemes and being reduced phonologically in the process.⁷¹

Iranian data indeed pattern in this way as some (but not all) auxiliaries are phonetically reduced while the heavy verb is not. This applies to the “transitive” preterite in Sogdian and Ossetic, where “hold” and “put” (if this is the right etymology) become fused with the main verb. Further reductions occur in the potential construction: in Sogdian, the main verb and “do” “coalesce into a single word, which is treated as a heavy stem (Sims-Williams 2007:379) while the non-finite forms (with *k²r* “deed”, cf. note 38) are not reduced; in the Xwarezmian potential, *KAR* is reduced to a particle, but preserved elsewhere.

Furthermore, *KAR* as a light verb is likewise phonetically reduced, but so is the heavy verb, confirming Butt’s argument of form identity of light and heavy verbs. The present stem of *KAR* (Old Persian *kunau-*, thence Middle and New Persian *kun-*, similarly in other Ir. languages) exhibits an irregular loss of the *r* vs. the corresponding Avestan (*kərənao-*) and Old Indic (*kṛṇo-*) forms; according to Hoffmann, this is an “allegro form”, i.e. the irregular

⁷¹ One example given by Butt (2010:65) is the Germanic preterite (cf. Section 5.4).

reduction is due to the verb being particularly frequent.⁷² Even more marked reductions are seen in Sogdian, which has <wn-> (besides <kwn->), and Khotanese, which only shows variants with irregular loss of the *k*- and hiatus-filling initial consonant, viz. *gan-*, *yan-*, *tan-*. These forms have been explained as arising from contexts where the verb was cliticised to another element,⁷³ but the forms are synchronically the only ones used.

While Iranian thus confirms Butt's position that light verbs are not part of the grammaticalisation cline shown by auxiliaries, the data appear to speak in favour of the development of light verbs and auxiliaries being parallel processes, thus agreeing with Bovern (2008, contra Butt 2010). Not only are the same verbs employed for both groups at least so far as the most common ones are concerned, but the stages of the development of the auxiliary and the light verb system seem to depend on each other.

Likewise contra Butt (2010:48f.), I argue that the semantic bleaching of the (ultimately) light verb is illustrable. In the case of *KAR*, the generalisation seems to be traceable in the data, proceeding from "make, do" to "perform, carry out" (thus in proto-CPs) or "succeed, achieve" (thus, as argued by Benveniste, the auxiliary in the potential construction) to a marker of prototypically transitive meaning definable by the features [+VOL, +INST, +AFF]. Further on, as the CP pattern becomes established, *KAR* develops to a marker of the feature [+INST].

As noted by Bovern (2008:168f.), it is of course not the light verbs in all individual CPs that undergo bleaching, but the verb loses properties and is analogically extended to more semantic contexts. This line of argument agrees well with the observation by several authors that CPs are not a uniform class, and there is "no clear-cut difference" between CPs and combinations with heavy verb (thus Thordarson 2009:78 for Ossetic). The degree of autonomy of the elements thus varies, because the CP formation is a productive and ongoing process: "there are all kinds of instances intermediary between a phrase such as *čāšm dūxtan* 'fix one's eyes upon [lit. sew the eye]', where the two members contribute equally to forming the meaning of the group, and a "compound verb" of the type *žārūb kardan* 'sweep [lit. do broom]', where the auxiliary [i.e., the light verb] only serves to transform a noun into an expression equivalent to a verb" (Telegdi 1951:322). This point is confirmed by alignment patterns in CPs in Ir. languages showing ergative structures. In Balochi, for instance, most CPs with a transitive light verb pattern ergatively even if their overall semantics is intransitive, but some have developed to an extent that they treat the subject as intransitive in spite of a transitive light verb.⁷⁴

5.5 Iranian also provides evidence for Bovern's (2008:163, 177) argument that gerunds (nominalised verbs) are an important source for CPs: *hamaranam* "battle" in (28) and *handarz* "advice" in (32) and (34) could be thus interpreted even if not a "nominalised verb" by the usual categories of Old and Middle Persian grammar.

Indeed, very much along these lines, I argue that participles play a role in the

⁷² The same would apply to irregular forms in Indic; see Hoffmann (1976:587f.) for more details and references.

⁷³ Sims-Williams (1983:44, 1990:290).

⁷⁴ Farrell (1995:232f.), Korn (2009:66f.).

emergence of the CP system as their nature of both verbal and nominal at the same time they lend them to a double interpretation. In the development of periphrastic verbal constructions such as the potential construction, they keep their verbal nature, but in a phrase such as “make / be remembered”, the former participle Khotanese *byāta* in (33) (Section 4.2.1) is liable to an interpretation of the nominal element in a complex predicate.

5.6 So far as the date of the development is concerned, it appears possible to establish at least some points. Given that *KAR* “do” is the most prominent and most widely used light verb in Iranian, the existence of this particular root appears to be presupposed. *KAR* acquires the meaning “do” only in Indo-Iranian, which would *a priori* exclude assumptions of complex predicates antedating Proto-Indo-Iranian.

[p. 53] Another *terminus post quem* is the formation of the causative-denominative in *-ēn-*, which seems to play a role in the rise of CPs in Persian. As this suffix is of Middle Iranian date, the emergence of complex predicates as a system appears to date to the Middle Persian period. This does not exclude the existence of individual combinations of noun and verb already at an earlier stage, but these appear to have been more idiomatic in nature.

Insofar as the age of individual CPs is concerned, it seems that Middle Persian *handarz kardan* “advise, instruct” can claim considerable age, since (34) would be the Middle Persian continuation of an Old Persian **handarzam KAR*, which is not attested, but implied by Aramaic *hndrz y^cbdwn* (32). Middle Persian *nām kardan* “call”, whose oldest datable occurrences are in the inscriptions of Kirdir (3rd century AD), one of which says (KNRb 27f.): “I am that Kirdēr whom Šābuhr King of Kings called Kirdēr the mowbed (...)", Skjærvø 1989:340).

As shown by (2), *nām kardan* being in competition with *nām(ēn)īdan* situates these instances probably on the level of the lexicon (see Section 2.2), potentially on the same level as German or Hungarian CPs such as shown in (3).

Contact languages provide some points for a *terminus post quem*. Aramaic texts from the 5th century BC showing noun + verb combinations probably copied from Iranian would appear to presuppose the existence in (a variety of) Old Persian of a number of idioms containing *KAR* “do” prominent enough to motivate their being copied. Syriac evidence would imply the same for Middle Persian. More light verbs only emerge in the next stage, evidenced by Armenian data. Armenian shows a groups of light verbs that include, beside “do”, “hold / have”, “hit” and “give” for the [+INST] pattern and “be” for [+AFF]. Agreeing with data from Middle Persian, Armenian CPs often coexist with a denominative of the same meaning.

The data appear to converge on suggesting that the Persian CP system emerged within the Middle Persian period, arising from a particular constellation of structures contributing to its development in combination with a group of more or less fixed expressions with *KAR*.

5.7 The approach presented here certainly does not claim to explain all of the Persian CP system. However, it seems that the line pursued here might offer an explanation to some inconsistencies observed by other authors. For instance, Folli et al. (2005:1377) note that verbs of movement do not entirely match their analysis. On the present approach, verbs of

movement are grouped with the verb “become” as light verbs marking [+AFF] and mirroring earlier middle voice inflection. Also, as noted by Megerdoomian 2001, the light verb *kardan* “do” does not always imply volition, and also includes unergatives such as “fly”. Her examples do imply the feature [+INST], though. Both points appear to agree rather well with the hypothesis presented here of [+INST] and [+AFF] being at the heart of the CP system.

A historical perspective is at the same time likely to suggest caution to any single explanation. Given the fact that the rise of CPs in Persian appears to be a gradual process that started in Middle Persian (with some prerequisites for the process arising much earlier), and that the process is an ongoing one, it is not to be expected that the data are uniform. The individual CPs will be (and are) of widely varying age, ranging from noun + verb combinations of considerable antiquity to those that are being formed by speakers today, yielding a very uneven pool of data. The investigation of the individual structures, and their interaction with each other, will certainly continue to be a fruitful field of research.

Acknowledgements

I owe an immense debt of gratitude to an anonymous reviewer for several rounds of detailed and committed feedback, which significantly contributed to the development of the argument of this paper. Sincere thanks also go to Thomas Jügel for commenting on a previous version.

References

- Abaev, V. I. 1964. *A Grammatical Sketch of Ossetic*. Mouton, The Hague
- Ahadi, Sh. 2001. *Verbergänzungen und zusammengesetzte Verben im Persischen. Eine valenztheoretische Analyse*. Reichert, Wiesbaden
- Amberber, M. 2010. The structure of the light verb construction. In: Amberber / Baker / Harvey, 291-318
- Amberber, M., Baker, B., Harvey, M. (Eds.) 2010. *Complex Predicates. Cross-linguistic Perspectives on Event Structure*. Cambridge University Press, Cambridge
- Axenov, S. 2006. *The Balochi Language of Turkmenistan*. Uppsala University, Uppsala
- Axvlediani, G. S. 1963. *Grammatika osetinskogo jazyka I: Fonetika i morfologija*. Naučno-issledovatel'skij institut pri sovete ministrov Severo-Osetinskoj ASSR, Ordzhonikidze
- Baker, B., Harvey, M. 2010. Complex predicate formation. In: Amberber / Baker / Harvey, 13-47
- Barjasteh Delforooz, B. 2007. Structure of Verbs in Brahui. In: Brunner, R., Laut, J. P., Reinkowski, M. (Eds.), XXX. *Deutscher Orientalistentag Freiburg, 24.-28. September 2007: Ausgewählte Vorträge*.
<http://webdoc.urz.uni-halle.de/dot2007/publikation.php>
- Barker, M. A., Mengal, A. Kh. 1969. *A Course in Baluchi*. McGill University Press, Montreal, 2 vol.
- Bashir, E. 2008. Some Transitional Features of Eastern Balochi: An Areal and Diachronic Perspective. In: Jahani / Korn / Titus, 45-82
- 2009. Wakhi. In: Windfuhr, 825-862
- Bedir Khan, E. Dj., Lescot, R. 1970. *Grammaire kurde (dialecte kurmanji)*. Maisonneuve, Paris

- Benveniste, É. 1929. Essai de grammaire sogdienne II: Morphologie, syntaxe et glossaire. Paul Geuthner, Paris
- 1954. Expression de « pouvoir » en iranien. Bulletin de la Société de Linguistique de Paris 50, 56-57 (= Benveniste, Émile 1979. Études sogdiennes. Reichert, Wiesbaden, 287-298) [p. 54]
- Bird, B. A. 2007. Aspects of Yaghnobi Grammar (unpublished MA thesis, University of Oregon), <http://yaghnobi.wordpress.com/yaghnobi-language/aspects-of-yaghnobi-grammar/>
- Blau, J. 1980. Manuel de kurde. Dialecte Sorani. Klincksieck, Paris
- Bowern, C. 2008. The diachrony of complex predicates. In: Bowern, C. (Ed.), The Diachrony of Complex Predication. Special issue of Diachronica 25/2, 161-185
- Brunner, Ch. 1977. A Syntax of Western Middle Iranian. Caravan Books, Delmar, NY
- Butt, M. 2010. The structure of the light verb construction. In: Amberber / Baker / Harvey, 48-78
- Canevascini, G. 1991. Medio-reflexive verbs in Khotanese. In: Emmerick, R., Weber, D. (Eds.), Corolla Iranica. Papers in honour of David Neil MacKenzie on the occasion of his 65th birthday on April 8th, 1991. Peter Lang, Frankfurt etc., 23-26
- Cheung, J. 2002. Studies in the Historical Development of the Ossetic Vocalism. Reichert, Wiesbaden
- Christol, A. 1990. Introduction à l'ossète: éléments de grammaire comparée. LALIES (Actes des sessions de linguistique et de littérature) 8, 7-50
- Ciancaglini, C. 2006. L'origine delle locuzione verbali con 'bd in siriano. In: Borbone, P. G., Mengozzi, A., Tosco, M. (Eds.), Loquentes linguas: Studi linguistici e orientali in onore di Fabrizio A. Pennacchietti. Harrassowitz, Wiesbaden, 173-184
- 2011. The Formation of the Periphrastic Verbs in Persian and Neighbouring Languages. In: Maggi, M., Orsatti, P. (Eds.), The Persian Language in History. Reichert, Wiesbaden, 3-21
- Dabir Moghaddam, M. 2006. Compound Verbs in Persian. In: Dabir Moghaddam, M.: Studies in Persian Linguistics: Selected articles. Iran University Press, Tehran, 65-114
- Durkin-Meisterernst, D. 2002. The uses of *baw-* in Parthian. In: Huyse, Ph. (Ed.), Iran. Questions et connaissances. Actes du IV^e congrès européen ... I. Association pour l'avancement des études iraniennes, Paris, 47-62
- 2009: Khwarezmian. In: Windfuhr, 336-376
- Emmerick, R. 1968. Saka Grammatical Studies. Oxford University Press, London etc.
- 1987. Auxiliaries in Khotanese. In: Harris, M., Ramat, P. (Eds.), Historical Development of Auxiliaries. Mouton de Gruyter, Berlin etc., 271-290
- 2009. Khotanese and Tumshuqese. In: Windfuhr, 377-425
- Farrell, T. 1990. Basic Balochi. An introductory course. Istituto Universitario Orientale, Naples
- 1995. Fading Ergativity? A Study of Ergativity in Balochi. In: Bennett, D., Bynon, Th., Hewitt, G. (Eds.), Subject, Voice and Ergativity: Selected Essays. SOAS, London, 218-243
- 2008. The Sweet Tongue: Metaphor in Balochi. In: Jahani / Korn / Titus, 101-138
- Foley, W. 2010. Events and serial verb constructions. In: Amberber / Baker / Harvey, 79-109
- Folli, R., Harley, H., Karimi, S. 2005. Determinants of event type in Persian complex predicates. Lingua 115, 1365-1401
- Fortson, B. 2010. Indo-European Language and Culture: An Introduction. Wiley-Blackwell, Chichester etc., 2nd ed.
- Fragner, B. 1999. Die "Persophonie". Regionalität, Identität und Sprachkontakt in der Geschichte Asiens. Das Arabische Buch, Berlin
- Fritz, S. 2009. Die „komplexen Verben“ des Neupersischen in Umgangssprache und Literatur – eine areallinguistische und kulturhistorische Betrachtung. Iranistik: Deutschsprachige Zeitschrift für

- iranistische Studien (Tehran) 6, 5-55
- Frye, R. 1966. The Persepolis Middle Persian Inscriptions from the Time of Shapur II. *Acta Orientalia* (Copenhagen) 30, 83-93
- Gershevitch, I. 1954. *A Grammar of Manichean Sogdian*. Oxford University Press, Oxford
- 1959. *The Avestan Hymn to Mithra, with an introduction, translation and commentary*. Cambridge University Press, Cambridge
- Gignoux, Ph. 1972. *Glossaire des inscriptions pehlevies et parthes*. Lund Humphries, London
- Gilbertson, G. W. 1923. *The Balochi Language. A grammar and manual*. Austin & Sons, Hertford
- Henning, W. B. 1934. *Das Verbum des Mittelpersischen der Turfanfragmente*. *Zeitschrift für Indologie und Iranistik* 9, 158-253 (=Henning, W.B. 1977. *Selected Papers I*. Brill, Leiden, 65-160)
- Hoffmann, K. 1955. *Altpersisch afuvāyā*. In: Krahe, H. (Ed.), *Corolla Linguistica. Festschrift Ferdinand Sommer (...)*. Wiesbaden: Harrassowitz, pp. 80-85 (here quoted from the reprint in: Hoffmann, K. 1975. *Aufsätze zur Indoiranistik I*. Reichert, Wiesbaden, 52-57)
- 1976. *Ved. karóti*. In: Hoffmann, K., *Aufsätze zur Indoiranistik II*. Reichert, Wiesbaden, 575-588
- Jahani, C. 1999. *Persian Influence on Some Verbal Constructions in Iranian Balochi*. *Studia Iranica* 28, 123-143
- Jahani, C., Korn, A. 2009. *Balochi*. In: *Windfuhr*, 634-692
- Jahani, C., Korn, A., Titus, P. (Eds.) 2008. *The Baloch and Others: Linguistic, historical and socio-political perspectives on pluralism in Balochistan*. Reichert, Wiesbaden
- Jespersen, O. 1965. *A Modern English Grammar on Historical Principles VI: Morphology*. George Allen and Unwin, London
- Jügel, Th. 2012. *Die Entwicklung der Ergativkonstruktion im Alt- und Mitteliranischen. Eine korpusbasierte Untersuchung zu Kasus, Kongruenz und Satzbau*. PhD thesis, University of Frankfurt a.M.
- Karimi-Doostan, Gh. 1997. *Light Verb Constructions in Persian*. Ph.D. thesis, Essex University.
- 2005. *Light Verbs and Structural Case*. *Lingua* 115, 1737-1756
- Kaufmann, I. 2007. *Middle Voice*. *Lingua* 117, 1677-1714
- Kazimirski, A. de Biberstein 1883. *Dialogues français-persans précédés d'un précis de la grammaire persane et suivis d'un vocabulaire français-persan*. Klincksieck, Paris
- Kellens, J. 1984. *Le Verbe Avestique*. Reichert, Wiesbaden
- Kemmer, S. 1993. *The middle voice*. Benjamins, Amsterdam
- Kent, R. 1953. *Old Persian. Grammar, Texts, Lexicon*. American Oriental Society, New Haven
- Klaiman, M. 1991. *Grammatical voice*. Cambridge University Press, Cambridge
- Klimčickij, S.I. 1937. *Jagnobsko-sogdijskie sootvetsvija*. *Zapiski instituta vostokovedenija Akademii Nauk* 6, 15-25
- Korn, A. 2009. *The Ergative System in Balochi from a Typological Perspective*. *Iranian Journal of Applied Language Studies* 1, 43-79
- Krahe, H., Meid, W. 1969. *Germanische Sprachwissenschaft II: Formenlehre*. de Gruyter, Berlin
- Kurdoev, K. K. 1978. *Grammatika kurdschego jazyka na materiale dialektov kurmandži i sorani*. Nauka, Moscow
- Lorenz, M. 1979. *Lehrbuch des Pashto (Afghanisch)*. VEB Verlag Enzyklopädie, Leipzig
- Lyons, J. 1968. *Introduction to Theoretical Linguistics*. Cambridge University Press, Cambridge
- MacKenzie, D. N. 1971. *Supplement*. In: Henning, W.B., *A Fragment of a Khwarezmian Dictionary*. Lund Humphries, London, 28-50

- Megerdooian, K. 2001. Event Structure and Complex Predicates in Persian. *Canadian Journal of Linguistics* 46, 97-125 [p. 55]
- 2012. The status of the nominal in Persian complex predicates. *Natural Language and Linguistic Theory* 30, 179-216
- Miller, W. 1903. Die Sprache der Osseten. Appendix to: Geiger, W., Kuhn, E. (Eds.), *Grundriß der iranischen Philologie I*. Trübner, Straßburg
- Mockler, E. 1877. A Grammar of the Baloochee Language, as it is spoken in Makrân (ancient Gedrosia), in the Persi-Arabic and Roman characters. Henry King, London
- Morgenstierne, G. 1927. An Etymological Vocabulary of Pashto. Dybwad, Oslo
- 2003. A New Etymological Vocabulary of Pashto. Reichert, Wiesbaden
- Næss, Å. 2007. Prototypical Transitivity. Benjamins, Amsterdam / Philadelphia
- Pantcheva, M. 2009. First Phase Syntax of Persian Complex Predicates: Argument Structure and Telicity. *Journal of South Asian Linguistics* 2 (online at <http://tiger.sprachwiss.uni-konstanz.de/~jsal/ojs/>)
- Paul, L. 1998. *Zazaki. Grammatik und Versuch einer Dialektologie*. Reichert, Wiesbaden
- 2009. *Zazaki*. In: *Windfuhr*, 545-586
- Pirejko, L. 1975. Kategorija zaloga. In: Vera S. Rastorgueva (Ed.), *Opyt istoriko-tipologičeskogo issledovanija iranskix jazykov II: Évoljucija grammatičeskix kategorij*. Nauka, Moscow, 302-336
- Rastorgueva, V. S., et al. 1971. *Giljanskij jazyk*. Nauka, Moscow
- Rastorgueva, V. S., Molčanova, E. K. 1981. *Srednepersidskij jazyk*. In: *Osnovy iranskogo jazykoznanija 2*. Nauka, Moscow, 6-146
- Reichelt, H. 1909. *Awestisches Elementarbuch*. Winter, Heidelberg
- 1931. Beiträge zur soghdischen Grammatik. In: Wüst, W. (Ed.), *Studia Indo-iranica*. Ehrengabe für Wilhelm Geiger. Harrassowitz, Leipzig, 248-260
- Robson, B., Tegey, H. 2009. Pashto. In: *Windfuhr*, 721-772
- Salemann, C. 1901. *Mittelpersisch*. In: Geiger, W., Kuhn, E. (Eds.), *Grundriss der iranischen Philologie I/1*. Trübner, Straßburg, 249-332
- Schmidt, R. L. 2003. Urdu. In: Cardona, G., Jain, Dh. (Eds.), *The Indo-Aryan Languages*. Routledge, London / New York, 286-350
- Schmitt, R. 1983. Iranisches Lehngut im Armenischen. *Revue des études arméniennes Nouvelle série*, XVII, pp. 73-112
- (ed.) 1989. *Compendium Linguarum Iranicarum*. Reichert, Wiesbaden
- 1991. *The Bisitun Inscriptions of Darius the Great: Old Persian Text*. SOAS, London
- 2009. *Die altpersischen Inschriften der Achaimeniden. Editio minor mit deutscher Übersetzung*. Reichert, Wiesbaden
- Sheintuch, G. 1976. Periphrastic verb formation in Persian: a dynamic process. In: Christie, W. (Ed.), *Current Progress in Historical Linguistics*. North-Holland Publishing Company, Amsterdam etc., 137-146
- Sims-Williams, N. 1983. Chotano-Sogdica. *Bulletin of the School of African and Oriental Studies* 46, 40-51
- 1989. Sogdian. In: Schmitt, 173-192
- 1990. Chotano-Sogdica II: Aspects of the Development of nominal Morphology in Khotanese and Sogdian. In: Gnoli, Gh., Panaino, A. (Eds.), *Proceedings of the First European Conference of Iranian Studies (...) 1: Old and Middle Iranian Studies*. Istituto Italiano del medio ed estremo oriente, Rome, 275-296

- 1994. The Middle Voice in Middle Iranian (handout of a paper presented at the "Symposion anlässlich des 25-jährigen Bestehens der Kommission für Iranistik an der Österreichischen Akademie der Wissenschaften, 4.-5. Nov. 1994", Vienna)
- 1997. The denominal suffix *-ant-* and the formation of the Khotanese transitive perfect. In: Lubotsky, A. (Ed.), *Sound Law and Analogy. Papers in honor of Robert S.P. Beekes on the occasion of his 60th birthday.* Rodopi, Amsterdam / Atlanta, 317-325
- 2007: The Sogdian potentialis. In: Macuch, M., Maggi, M., Sundermann, W. (Eds.), *Iranian Languages and Texts from Iran and Turfan. Ronald E. Emmerick Memorial Volume.* Harrassowitz, Wiesbaden, 377-386
- Skjærvø, P. O. 1989. Verbal Ideograms and the Imperfect in Middle Persian. In: de Fouchécour, Ch.-H., Gignoux, Ph. (Eds.), *Études irano-aryennes offertes à Gilbert Lazard.* Association pour l'avancement des études iraniennes, Paris, 333-354
- 2009. Old Iranian. In: Windfuhr, 43-195
- 2009a. Middle West Iranian. In: Windfuhr, 196-278
- Sundermann, W. 1981. *Mitteliranische manichäische Texte kirchengeschichtlichen Inhalts.* Akademie der Wissenschaften der DDR, Berlin
- 1989. *Mittelpersisch.* In: Schmitt, 138-164
- Telegdi Zs. 1951. Nature et fonction des périphrases dites 'verbes composés' en Persan. *Acta Orientalia Academiae Scientiarum Hungaricae* 1, 315-338 (repr. in: *Telegdi Zs. 2006, Opera Omnia I.* Akadémiai Kiadó, Budapest, 123-144)
- Thordarson, F. 1989. Ossetic. In: Schmitt, 456-479
- 2009. *Ossetic Grammatical Studies.* Österreichische Akademie der Wissenschaften, Vienna
- Wendtland, A. 2011. The Emergence and Development of the Sogdian Perfect. In: Korn, A., Haig, G., Karimi, S., Samvelian, P. (Eds.), *Topics in Iranian Linguistics.* Reichert, Wiesbaden, 39-52
- Windfuhr, G. 1979. *Persian Grammar. History and State of its Study.* Mouton, The Hague etc.
- (ed.) 2009. *Iranian Languages.* RoutledgeCurzon, London
- Xromov, A. L. 1972. *Jagnobskij jazyk.* Nauka, Moscow
- Yoshida Y. 2009. Sogdian. In: Windfuhr, 279-335
- Zaršenās, Z. 2002. Šavāhed-ī az sāxt-e emkānī dar zabān-e balūčī. In: Režāi Bāgbīdī, Ḥ. (Ed.), *Maǰmū'e-ye maqālāt-e naxostīn ham-andīše-ye gūyeš-šenāsī-ye īrān.* Tehran 1381 h.š, 259-270