Predicative multi-word expressions in Persian

Jens Fleischhauer
Department of General Linguistics
Heinrich-Heine-Universität, Düsseldorf, Germany
fleischhauer@phil.uni-duesseldorf.de

Abstract

Persian, like many other Asian languages, licenses the use of bare nouns in object position. Such sequences are often treated as multi-word expressions (compound verbs/light verb constructions, and pseudo-incorporation constructions). In the paper, I argue against a uniform treatment of all 'bare noun + verb' sequences in contemporary Persian. The paper presents criteria which allow to distinguish light verb constructions from other superficially similarly looking predicational construction types.

1 Introduction: Predicative construction types

Persian, an Iranian language of Western Asia, has a rather small set of lexically full verbs. Mohammad & Karimi (1992, 195) mention a number of around 115 lexically full verbs; others, for example Samvelian & Faghiri (2016, 212) and Samvelian (2018, 256), mention somewhat higher numbers (about 250) but state that only around half of them are still in use. The lack of full verbs is compensated by the use of light verb constructions (sometimes also called 'compound verbs'). Light verb constructions (LVCs) consist of a semantically reduced verb and a non-verbal element (NVE). The NVE is prototypically an NP (cf. the LVC *sedâ dâdan* 'to produce a sound' in (1)).[1]

(1) *Sag sedâ dâd.*
dog sound give.PST
'The dog made a sound.'

Light verb constructions are multi-word expressions since they form a predicational unit consisting of (at least) two words. LVCs are fixed expressions since the set of light verbs is rather restricted (according to Family (2006, 8), around 20 Persian full verbs exhibit light and heavy uses). Furthermore, the combination of a light verb with a specific type of NVE is not fully predictable. Nevertheless, LVCs are semantically compositional as they show properties very similar to those found with idiomatically combining expressions (in the sense of Nunberg et al. 1994); LVCs license internal modification and form families (Family, 2011; Samvelian, 2012; Samvelian & Faghiri, 2014, 2016; Fleischhauer et al., 2019; Fleischhauer & Gamerschlag, 2019; Fleischhauer & Neisani, 2020).

Light verbs are formally identical to the lexically full verbs as the examples in (2) show. The light use of *xordan* 'eat' is illustrated in (2a), there it heads the LVC *qose xordan* 'to worry about' (lit. 'concern eat'). The example in (2b) illustrates the heavy use of *xordan* which denotes an event of eating food.

(2) a. *hâmiše qosey-e fârda-ra*
always concern-EZ tomorrow-ACC
mi-xor-âd.
IMPF-eat-3SG
'She always worries about the future.'
(Family, 2006, 85)

[1] Glossing of the examples follows the Leipzig glossing rules; the following abbreviations are used: ABS: absolutive, ACC: accusative, CL: classifier, DEM: demonstrative, EMPH: emphatic, ERG: ergative, EZ: ezâfe, IMPF: imperfective aspect, INDEF: indefinite, NEG: negation, PST: past tense, PL: plural, SG: singular, SUB: subjunctive.

b. *Bâčče-ha qazâ* râ *xord-and*.
child-PL food ACC eat.PST-3PL
'The children ate the food.'

There is some debate whether all sequences of bare noun plus verb should be conceived as instances of the same type of complex predicate or not. Some authors (e.g., Ghomeshi & Massam 1994; Vahedi-Langrudi 1996; Mahmoodi-Bakhtiari 2018, 295) treat the 'N + V' sequences in (3) equally as 'compound verbs.' Others (e.g., Mohammad & Karimi, 1992; Lazard, 1992; Nemati, 2010; Megerdoomian, 2012; Modaressi, 2014) argue that the two sequences in (3) look superficially similar but exemplify different types of constructions. The example in (3a) is analyzed as an instance of pseudo-incorporation (e.g. Nemati 2010; Modaressi 2014, 2015); the one in (3b) is an LVC.

(3) a. *Bâčče-ha qazâ xord-and*.
child-PL food eat.PST-3PL
'The children ate (food).'
b. *Doxtar dârad jiq mi-zan-ad*.
girl has scream IMPF-hit-3SG
'The girl is screaming.'

Thus, Persian poses the problem of distinguishing between at least three predicate construction types: (i) regular predicate-argument constructions (2b), (ii) light verb constructions (1), (2a) and (3b), and (iii) pseudo-incorporation constructions (3a). I define a predicational construction type as a specific morphosyntactic construction which realizes the sentence predicate. Most importantly, the set of light verbs is partially overlapping with the set of heavy verbs which show pseudo-incorporation of bare nouns in object position. An example is the verb *xordan* 'to eat', which has already been illustrated in the examples in (2a) and (3a) above.

Irrespective of the question whether one analyzes examples like in (3a) as multi-word expression or not (see, e.g., Hüning & Schlücker 2015), one needs criteria to distinguish between LVCs on the one hand and pseudo-incorporation constructions (PICs) on the hand. I will argue that such a set of criteria will also allow to distinguish between LVCs and regular predicate-argument constructions. Although the current analysis focuses on Persian only, the question discussed in the paper is highly relevant for a larger number of Asian (but also non-Asian) languages (e.g. Turkish, Hindi/Urdu, Kurdish).

2 Pseudo-Incorporation

The term 'pseudo-incorporation' goes back to Massam's (2001) analysis of Niuean, which is a Malayo-Polynesian language. Niuean exhibits a grammatical phenomenon which is reminiscent of the nominal incorporation attested, for example, in many American languages (e.g., Mohawk or Tiwa, see, Mithun 1986). In Niuean, a nominal element is usually preceded by a case marker, as indicated in (4a). The absence of a case marker preceding *ika* 'fish' (4b) results in a change from a transitive case frame (ergative case for the subject, absolutive case for the object) to an intransitive one (absolutive case for the subject).

(4) a. *Takafaga tūmau nī e ia e*
hunt always EMPH ERG he ABS
tau ika.
PL fish
'He is always fishing.'
b. *Takafaga ika tūmau nī a ia*.
hunt fish always EMPH ABS he
'He is always fishing.'
(Massam, 2001, 157)

A crucial aspect of the 'V + N' sequence *takafaga ika* 'hunt fish' in (4b) is that the two do not form a morphological word (for details, the reader is referred to the original discussion in Massam 2001). Subsequent work has shown that pseudo-incorporation is widespread among the world's languages. The literature on (pseudo)-incorporation has identified a number of stable semantic properties which are cross-linguistically attested in (pseudo)-incorporation this issue, see Borik & Gehrke 2015). Pseudo-incorporated nouns tend to be non-referential and show the following properties: (i) they have obligatory narrow scope with respect to scope bearing elements, e.g., negation; (ii) they are number neutral; (iii) they are discourse opaque; and (iv) they show restrictions with respect to modifiability. One has to mention that there is some cross-linguistic variance with respect to these properties; especially the property of discourse opacity is somewhat relaxed in some lan-

guages (see Farkas & de Swart 2003 on Hungarian). I illustrate these properties by the use of Persian language data.

Starting with the first property, the bare noun *gorbeh* 'cat' in (5a) has narrow scope with respect to the negation operator. The only interpretation of the sentence is that the subject referent did not see any cat. The non-bare, i.e., case marked, noun in (5b) has wide scope with respect to the negation operator. The sentence means that there is a particular cat which the subject referent did not see.

(5) a. *Gorbeh na-did-âm.*
cat NEG-see.PST-1SG
'I didn't see any cat.' [¬ > ∃]
b. *Gorbeh-râ na-did-âm.*
cat-ACC NEG-see.PST-1SG
'I didn't see the cat.' [∃ > ¬]

Number neutrality is illustrated by the example in (6a). The noun *gorbeh* 'cat' is used without a plural marker but licenses a singular as well as plural interpretation. With respect to non-bare nouns, number interpretation depends on number marking. If the noun neither bears plural marking nor is preceded by a number word, it only licenses a singular reading (6b).

(6) a. *Gorbeh did-âm.*
cat see.PST-1SG
'I saw (a) cat/cats.'
b. *Gorbeh-râ did-âm. #Xeili ziba*
cat-ACC see.PST-1SG very pretty
bood-ând.
be.PST-3PL
'I saw the cat. #They were very pretty.'

Bare nouns are non-referential and therefore do not introduce discourse referents. The bare noun *šer* 'poem' in (7a) cannot serve as the antecedent of a (null) anaphora.[2] In non-bare use, the noun introduces a discourse referent and cannot be picked up anaphorically (7b).

(7) a. *Ali bâyad šer be-xân-ad. #Ân*
Ali must poem SUB-read-3SG DEM
(*šer*) *tavasote yek šâer-e arab*
POEM by INDEF poet-EZ Arabic
sorude šod.
written become
'Ali must read a poem. That poem was written by an Arabic poet.'
b. *Ali bâyad šer-i bexânad.*
Ali must poem-INDEF SUB-read-3SG
Ân (*šer*) *tavasote yek šâer-e*
DEM poem by INDEF poet-EZ
arab sorude šod.
Arabic written become
'Ali must read a [specific] poem. That poem was written by an Arabic poet.'

Finally, pseudo-incorporated nouns are restricted with respect to modification. Attributively used adjectives require a linking element — called 'ezâfe' -– which is an affix placed between the modified noun and its modifier. As (8a) shows, the bare noun does not license the adjective *ziba* 'beautiful' as an attributive modifier. Modification is, as shown in (8b), restricted to kind-level modifiers (a similar restriction is mentioned by Espinal & McNally 2011 for Spanish and Catalan).

(8) a. **Ketab-e ziba nevešt-âm.*
book-EZ beautiful write.PST-1SG
'I write (a) beautiful book/books.'
b. *Mân ketab-e ghesseh mi-xâr-âm.*
I book-EZ story IMPF-buy-1SG
'I buy story books.'
(Modaressi, 2014, 23)

The properties associated with pseudo-incorporation are only found with bare nouns in object position. Other types of bare nouns, especially those figuring as the subject argument, do not show these properties.

3 Semantic differences between LVCs and pseudo-incorporation constructions

The current section aims at demonstrating that the semantic function of the verb is different in LVCs and PICs. Light verbs are semantically reduced compared to their heavy verb use. They do not have

[2]Modaressi (2014, 2015) as well as Krifka & Modarresi (2016) show that Persian bare nouns are not fully discourse opaque but are discourse translucent, following the terminology of Farkas & de Swart (2003). For a discussion of this issue, the reader is referred to the mentioned literature.

full predicational content and therefore do not denote an event of their own (e.g., Butt & Geuder, 2001, 356). Rather, the denoted eventuality is mainly determined by the NVE. This becomes clear from Fillmore et al.'s (2003) discussion of the differences between the English verb *decide* and the LVC *make a decision*. They write that both sentences in (9) "report on the same event, that of deciding something" (Fillmore et al., 2003, 244).[3] Although the LVC is headed by the light verb *make*, the authors state that sentence (9b) is "not about an event of making." This is tantamount to saying that the LVC denotes a different situation-type -– or event-type -– than the one denoted by the heavy correspondent of its verbal head.

(9) a. *The committee decided to convene again next month.*
b. *The committee made a decision to convene again next month.*

Building on the above-mentioned idea, I propose the working definition of a light verb construction presented in (10).

(10) A light verb construction is a complex predicate consisting of a semantically light verb and a non-verbal element. The situation type denoted by the light verb construction is not a subtype of the situation type denoted by the heavy verb but is dependent on the NVE.

The basic idea is that the light verb construction denotes a different type of situation than the heavy verb. Since this is a crucial part of the argumentation, I like to illustrate the definition by use of the Persian examples in (11).

(11) a. *Ân mard be ân zan yek*
DEM man to DEM woman INDEF
ketâb dâd.
book give.PST
'The man gave a book to the woman.'
b. *Sag sedâ dâd.*
dog sound give.PST
'The dog made a sound.'

[3]Fillmore et al. (2003) do not use the term 'light verb' but speak of 'support verbs.'

In (11a), *dâdan* is used as a heavy verb and denotes a giving-situation, which requires a special relation between an agent (the giver), a theme (the given), and a recipient. In this type of situation, the referent of the theme is transferred from the giver to the recipient. The LVC *sedâ dâdan* 'produce a sound', in (11b), denotes a situation of sound emission, which involves an emitter and an emittee (the emitted sound). A sound emission-situation is not a specific subtype of a giving-situation. That the two constructions denote different situation types is evidenced by the fact that only the example in (11a) allows adding 'and she is still in possession of it.' Thus, the working definition captures the basic idea that the main predicational content of an LVC is contributed by the non-light element, whereas the light verb merely adds information to the event predication (e.g., Butt & Geuder, 2001, 2003).

The definition in (10) allows us to distinguish between LVCs on the one hand and PICs on the other. As mentioned above, in the case of an LVC like *sedâ dâdan* 'produce a sound', the denoted situation type is not determined by the verb but by the NVE. In the case of a PIC like in (12), the verb determines the denoted situation. *Gorbeh didan* 'see a cat(s)' is a specific subtype of a seeing situation, i.e., it is a seeing of cats rather than of some other stimulus. The pseudo-incorporated noun further specifies the situation type denoted by the verb.

(12) *Gorbeh did-âm.*
cat see.PST-1SG
'I saw (a) cat/cats.'

This brief discussion gives rise to a first distinguishing property of light verb constructions and pseudo-incorporation constructions:

(13) LVCs differ from PICs with respect to the lexical element(s) determining the denoted situation type.

The contrast with respect to the element(s) determining the denoted situation type is a direct consequence of a difference regarding the status of the verb in the two types of complex predicates. The verbal head of an LVC is a light verb, whereas it is a heavy verb in the case of a PIC. Evidence for this fact is gained from the interpretation of the examples

discussed above. *Sedâ dâdan*, as already discussed in some detail, does not mean 'to give a sound to someone'. Thus, *dâdan* does not contribute its full lexical content. In the case of the PIC *gorbeh didan* 'cat see', the verb contributes its full lexical content. Only if a verb is used as a heavy verb and contributes its full lexical content is it able to determine the denoted situation type. I summarize this as a second distinguishing feature between the two types of constructions:

(14) The verbal head of an LVC is a light verb; a PIC is headed by a heavy verb.

After having present semantic differences between LVCs and PICs, I turn next to the discussion of the role bare nouns play in the two predicational construction types.

4 Bare nouns as NVEs

The semantic properties of pseudo-incorporated nouns are usually only found with nouns showing "some degree of bareness" (Borik & Gehrke, 2015, 12). In some languages, pseudo-incorporation is restricted to nouns without any functional morphology (e.g., number, case, in/definiteness marking), while other languages show weaker restrictions. Hungarian (Farkas & de Swart, 2003) and Greek (Gehrke & Lekakou, 2013) license accusative case marking on pseudo-incorporated nouns, whereas Hindi allows plural marking (Dayal, 2011). The discussion in Section 2 revealed that Persian restricts pseudo-incorporation to bare nouns.

The current section aims at investigating two different albeit related questions: First, do bare nouns used as NVEs of light verb constructions show the same semantic properties than pseudo-incorporated nouns? Second, is the nominal element within an NVE necessarily bare or does is license functional morphology?

4.1 The interpretation of bare noun NVEs

In section 2, it was shown that pseudo-incorporated nouns show a number of recurrent properties: they have narrow scope with respect to scope bearing elements, they are number neutral, they are discourse translucent, and, finally, they only license kind-level modifiers. The crucial question to be answered in the current section is whether bare nouns used as the NVE of a light verb construction show the same properties.

For the purpose of illustration, I will use the LVC *sedâ dâdan* 'produce a sound'. As the example in (15) shows, the bare noun *sedâ* does not introduce a discourse referent. Furthermore, the noun is interpreted as number neutral; the LVC either refers to situations of emitting a single sound or of emitting a number of (different or non-different) sounds.

(15) *#Âbgarmkon sedâ dâd. Ân (sedâ)*
boiler sound gave DEM sound
boland bud.
loud be.PST
Intended: 'The boiler produced (a) sound(s). It was loud.'
(Fleischhauer & Neisani, 2020, 13)

The bare noun NVE also has narrow scope with respect to negation. It is understood that the boiler did not produce any sound rather than that there is a particular sound which it did not produce.

(16) *Âbgarmkon sedâ na-dâd.*
boiler sound NEG-give.PST
'The boiler did not produce any sound.'

In contrast to pseudo-incorporated nouns, bare noun NVEs show fewer restrictions with respect to modification. As (17) shows, the bare noun *sedâ* licenses modification by the adjective *boland* 'loud' which is not a kind-level modifier.

(17) *Sedâ-ye boland dâdan nešân az*
sound-EZ loud give sign from
qodrat nist.
strength NEG.be.3SG
'Producing a loud sound/loud sounds is not a sign of strength.'
(Fleischhauer & Neisani, 2020, 13)

Bare noun NVEs indeed share a number of properties with pseudo-incorporated nouns. One might take this as evidence that there is no real distinction between LVCs on the one hand and PICs on the other. Contrary to this assumption, LVCs and PICs show a number of differences, for example, regarding restrictions on nominal morphology.

4.2 Morphosyntactic properties of pseudo-incorporated nouns and NVEs

With respect to light verb constructions, the question is whether LVC-formation is – like pseudo-incorporation – similarly restricted to nominal elements showing some degree of bareness. Persian has nominal morphology for the expression of indefiniteness, number as well as case. The three categories are briefly discussed subsequently.

Indefiniteness marking

Persian has different grammatical means for expressing indefiniteness: the indefinite article *yek* – which is identical to the numeral 'one' – and the phrasal suffix *-i*. The two markers have an overlapping but not identical distribution and can also be used in combination (Ghomeshi 2003, 65, Paul 2008, 322, Fleischhauer & Neisani 2020, 11). Within the limits of the current paper, I cannot present a detailed discussion of the similarities and differences of the indefiniteness markers. For the current discussion, it seems sufficient to say that *-i* signals specificity, whereas *yek* does not. Only *-i* but not *yek* can be used in referentially opaque contexts, like in (18a). In the example in (18b), either *yek* or *-i* but also both together can be used.

(18) a. *Ahmad mi-xâst bâ* (**yek*)
Ahmad IMPF-want.PST with INDEF
zan-e puldâr-i ezdevâj
woman-EZ rich-INDEF marry
kon-ad ammâ na-tavânest
do-3SG but NEG-could
kas-i-râ peidâ kon-ad.
one-INDEF-ACC find do-3SG
'Ahmad wanted to marry a rich woman but could not find one.'

b. *Ahmad mi-xâst* bâ (*yek*)
Ahmad IMPF-want.PST with INDEF
zan-e puldâr(*-i*) *ezdevâj*
woman-EZ rich-INDEF marry
kon-ad ammâ u tark-aš kard.
do-3SG but she leave-3SG did
'Ahmad wanted to marry a rich woman but she left him.'
(slightly adapted from Fleischhauer & Neisani 2020, 11f.)

Nouns marked for indefiniteness – either by *yek* or *-i* – do not show the properties of pseudo-incorporated nouns. Rather, despite the fact that *-i* can be used in referentially opaque contexts, indefinite nouns are discourse transparent (18b). Additionally, nouns marked for indefiniteness receive a number specific interpretation, i.e., they are not number neutral. This is evidenced in (19): *yek* as well as *-i* enforce a singular interpretation of the noun *medad* 'pencil'. Thus, a specification on the number of pencils cannot be added to the sentences in (19).

(19) a. *Yek medad avord-âm*,
INDEF pencil bring.PST-1SG
#yek-i bâraye khod-âm va
one-INDEF for self-1SG and
do-ta bâraye Leila.
two-CL for Leila
'I brought a pencil (one for me and two for Leila).'

b. *Medad-i avord-âm*,
pencil-INDEF bring.PST-1SG
#yek-i bâraye khod-âm va
one-INDEF for self-1SG and
do-ta bâraye Leila.
two-CL for Leila
'I brought a [specific] pencil (one for me and two for Leila).'
(based on Modaressi 2014, 24)

The nominal element within an NVE can be marked for indefiniteness, as the example in (20) shows, although the eventive noun *sedâ* 'sound' receives a referentially specific interpretation and introduces a discourse referent. This is evidenced by the fact that the referent introduced by *sedâ* can be anaphorically picked up. Thus, although *sedâ* is used referentially, it still forms a complex predicate with the verb *dâdan*. Irrespective of whether *sedâ* is marked for indefiniteness or not, the combination of *sedâ* with *dâdan* is interpreted as 'produce (a) sound(s)' rather than 'give someone a sound'. Thus, *dâdan* is still used as a light verb in (20) rather than as a heavy verb. This demonstrates that there is no relevant difference between the indefiniteness marking of NVEs and that of 'regular' nouns in argument position.

Nominal morphology	Pseudo-incorporation	Light verb construction
case	no	yes
indefiniteness	no	yes
number	no	yes

Table 1: Nominal morphology in Persian complex predicates.

(20) *Âbgarmkon sedâ-i dâd. Ân*
boiler sound-INDEF gave DEM
(*sedâ*) *boland bud.*
sound loud be.PST
'The boiler produced a [specific] sound. That (sound) was loud.'
(Fleischhauer & Neisani, 2020, 13)

Number marking

Persian has a binary number system distinguishing between an unmarked singular and a morphologically expressed plural. The plural marker *-hâ* is optional in contexts in which number is already expressed by other means, e.g., number words. The example in (21) shows plural marking of the NVE *sedâ*; the interpretation of the example is that the subject referent produces a number of (different) sounds.

(21) *In mâšin šab-hâ sedâ-i-hâ*
DEM car night-PL sound-INDEF-PL
mi-dah-ad.
IMPF-give-3SG
'This car produces some [specific] sounds at night.'
(Fleischhauer & Neisani, 2020, 12)

Case marking

Persian has a binary case system: it possesses a morphologically unmarked nominative case and the phrasal case affix *-râ* which marks accusative case. The language displays definiteness-based differential object marking, restricting accusative case marking to nouns that have a referentially specific interpretation (see, e.g., Bossong 1985; Lazard 1992, and Ghomeshi 1997). Since NVEs take the specificity marker *-i*, it is not surprising that they also license accusative case marking. An example taken from Karimi-Doostan (2011, 89) is shown in (22).[4]

[4]For more data on the case marking of the non-verbal element of Persian LVCs, see, e.g., Samvelian & Faghiri (2014, 51) and Karimi-Doostan (1997/2012, 203ff.).

The LVC under discussion is *râhnamâʔi kardan* 'advice/give advice'; the NVE *râhnamâʔi* 'advice' bears accusative case marking and is separated from the light verb by the indirect object *be Sasan* 'to Sasan'.

(22) *Ali in râhnamâʔi-râ be Sasan kard.*
Ali DEM advice-ACC to Sasan do.PST
'Ali gave Sasan this advice.'

Interim summary

The nominal elements used within the two types of complex predicates have different morphosyntactic properties: pseudo-incorporated nouns have a higher degree of bareness than NVEs. Whereas pseudo-incorporated nouns have to be bare, NVEs do not carry restrictions with respect to number, indefiniteness, or case marking. This does not mean that any NVE licenses all types of functional morphology; accusative case marking, for example, is restricted to NVEs which are realized as the light verb's direct object. So far there has been no systematic investigation of which NVEs are realized as a direct object.

The morphological properties of the two predicative construction types are summarized in table 1. Pseudo-incorporation but not LVC-formation is restricted to non-referential nouns, i.e., nouns which are neither marked for case nor for indefiniteness or number. Thus, nominal morphology allows us to distinguish LVC-formation from pseudo-incorporation but does not provide clear-cut criteria for identifying NVEs. Bare noun NVEs superficially look like pseudo-incorporated verb complements, while non-bare noun NVEs superficially look like non-pseudo-incorporated verb complements.

5 Conclusion

The starting point of the current paper was the question whether all instances of 'bare noun + verb'

Predicational construction type	**Noun**	**Verb**
pseudo-incorporation construction	semantics: non-referential object argument morphosyntax: bare direct object noun	heavy verb determines situation type
regular (transitive) predicate-argument construction	semantics: referential object argument morphosyntax: non-bare direct object noun	heavy verb determines situation type
light verb construction	semantics: *no restrictions on referentiality* morphosyntax: *no restriction* determines situation type	light verb

Table 2: Summary of the semantic and morphosyntactic properties of Persian predicational construction types.

exemplify the same type of complex predicate or not. The current paper argues against this view and presents evidence for a distinction between LVCs and PICs. LVCs can be distinguished from PICs by a number of properties. First, the verbal head of a light verb construction is semantically light, i.e., it does not contribute its full lexical content. The verbal head of a PIC, on the other hand, is a heavy verb.

Second, the denoted situation type of an LVC is determined by the NVE but not by the light verb. This point is closely related to the first one mentioned above since light verbs are defective event predicates. In the case of pseudo-incorporation, the denoted situation type is a subtype of the situation type denoted by the verbal head, i.e., food-eating is a subtype of eating but sound emission (literally 'give a sound') is not a subtype of giving-situations.

Third, nominal morphology does not block LVC-formation. Rather, NVEs are basically compatible with all types of nominal morphology. Most crucially, NVEs license case and number as well as indefiniteness marking, which is compatible with the fact that LVC-formation is not restricted to non-referential NVEs. Case as well as specificity marking block pseudo-incorporation.

The morphosyntactic as well as semantic differences between the three basic predicational construction types discussed in the current paper are summarized in table 2.

Among the questions which need to be addressed in future work is the following: Are there also syntactic differences between LVCs and PICs? Such differences are expected given that pseudo-incorporation seems to be restricted to bare nouns in immediately preverbal position. On the basis of the criteria present in the current paper, a corpus-based study on the syntactic behavior of LVCs and PICs is planned.

The identification of semantic, morphosyntactic and syntactic properties of different types of MWEs will hopefully a better identification of these expression in language corpora.

Acknowledgments

The research was carried out as part of the research project 'Funktionsverbgefüge: Familien & Komposition' ('Light verb constructions: Families & composition'; HE 8721/1-1) funded by the Deutsche Forschungsgemeinschaft (DFG). I like to thank Mozhgan Neisani for help with the language data.

References

Borik, Olga & Berit Gehrke. 2015. An Introduction to the Syntax and Semantics of Pseudo-Incorporation. In Olga Borik & Berit Gehrke (eds.), *The Syntax and Semantics of Pseudo-Incorporation*, 1–43. Leiden/Boston: Brill.

Bossong, Georg. 1985. *Empirische Universalienforschung: Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Gunter Narr.

Butt, Miriam & Wilhelm Geuder. 2001. On the (Semi)Lexical Status of Light Verbs. In Norbert Corver & Henk van Riemsdijk (eds.), *Semilexical Categories: On the content of function words and*

the function of content words, 323–370. Berlin: Mouton.

Butt, Miriam & Wilhelm Geuder. 2003. Light verbs in Urdu and grammaticalization. In Regine Eckardt, Klaus von Heusinger & Christoph Schwarze (eds.), *Words in Time*, 295–350. Berlin/New York: Mouton de Gruyter.

Dayal, Veneeta. 2011. Hindi pseudo-incorporation. *Natural Language & Linguistic Theory* 29(1). 123–167.

Espinal, M. Theresa & Louise McNally. 2011. Bare nominals and incorporating verbs in Spanish and Catalan. *Journal of Linguistics* 47. 87–128.

Family, Neiloufar. 2006. *Explorations of Semantic Space: The Case of Light Verb Constructions in Persian*. Paris: Ecole des Hautes Etudes en Science Sociales dissertation.

Family, Neiloufar. 2011. Verbal islands in Persian. *Folia Linguistica* 45(1). 1–30.

Farkas, D. & H. de Swart. 2003. *The Semantics of Incorporation: From Argument Structure to Discourse Transparency*. Stanford: CSLI Publications.

Fillmore, Charles, Christopher Johnson & Miriam Petruck. 2003. Background to FrameNet. *International Journal of Lexicography* 16(3). 235–250.

Fleischhauer, Jens & Thomas Gamerschlag. 2019. Deriving the meaning of light verb constructions – a frame account of german *stehen* 'stand'. In Constanze Juchem-Grundmann, Michael Pleyer & Monika Pleyer (eds.), *Yearbook of the German Cognitive Linguistics Association, Vol. 7*, 137–156. Berlin/Boston: Mouton de Gruyter.

Fleischhauer, Jens, Thomas Gamerschlag, Laura Kallmeyer & Simon Petitjean. 2019. Towards a compositional analysis of German light verb constructions (LVCs) combining Lexicalized Tree Adjoining Grammar (LTAG) with frame semantics. In *Proceedings of the 13th International Conference on Computational Semantics - Long Papers*, 79–90. Gothenburg, Sweden: Association for Computational Linguistics. https://www.aclweb.org/anthology/W19-0407.

Fleischhauer, Jens & Mozhgan Neisani. 2020. Adverbial and attributive modification of Persian separable light verb constructions. *Journal of Linguistics* 56. 45–85.

Gehrke, Berit & Marika Lekakou. 2013. How to miss your preposition. *Studies in Greek Linguistics* 33. 92–106.

Ghomeshi, Jila. 1997. Topics in Persian VPs. *Lingua* 102(2–3). 133–167.

Ghomeshi, Jila. 2003. Plural marking, indefiniteness, and the noun phrase. *Studia Linguistica* 57(2). 47–74.

Ghomeshi, Jila & Diane Massam. 1994. Lexical/syntactic relations without projections. *Linguistic Analysis* 24 (3–4). 175–217.

Hüning, Matthias & Barbara Schlücker. 2015. Multi-word expressions. In Peter Müller, Ingeborg Ohnheiser, Susan Olsen & Franz Rainer (eds.), *Word formation: An International Handbook of the Languages of Europe, Vol. 1*, 450–467. Berlin/Boston: De Gruyter Mouton.

Karimi-Doostan, Gholamhossein. 1997/2012. *Light Verb Constructions in Persian*. Saarbrücken: Lambert Academic Publishing.

Karimi-Doostan, Gholamhossein. 2011. Separability of Light Verb Constructions in Persian. *Studia Linguistica* 65(1). 70–95.

Krifka, Manfred & Fereshteh Modarresi. 2016. Number neutrality and anaphoric update of pseudo-incorporated nominals in Persian (and weak definites in English). In Mary Moroney, Carol-Rose Little, Jacob Collard & Dan Burgdorf (eds.), *Proceedings of Semantics and Linguistic Theory (SALT 26)*, 874–891. Washington DC: Linguistic Society of America.

Lazard, Gilbert. 1992. *A grammar of contemporary Persian*. Costa Mesa, CA/New York: Mazda Publishers.

Mahmoodi-Bakhtiari, B. 2018. Morphology. In A. Sedighi & P. Shabani-Jadidi (eds.), *The Oxford Handbook of Persian Linguistics*, 273–299. Oxford: Oxford University Press.

Massam, Diane. 2001. Pseudo Noun Incorporation in Niuean. *Natural Language & Linguistic Theory* 19. 153–197.

Megerdoomian, Karime. 2012. The status of the nominal in Persian complex predicates. *Natural Language & Linguistic Theory* 30. 179–216.

Mithun, Marianne. 1986. On the nature of noun incorporation. *Language* 62(1). 32–37.

Modaressi, F. 2014. *Bare nouns in Persian: Interpretation, Grammar and Prosody*. Ottawa/Berlin: University of Ottawa & Humboldt-Universität zu Berlin dissertation.

Modaressi, F. 2015. Discourse Properties of Bare Noun Objects. In Olga Borik & Berit Gehrke (eds.), *The Syntax and Semantics of Pseudo-Incorporation*, 189–221. Leiden/Boston: Brill.

Mohammad, Jan & Simin Karimi. 1992. Light verbs are taking over: Complex predicates in Persian. In J. A. Nevis & V. Samiian (eds.), *Proceedings of the Western Conference on Linguistics, vol. 5*, 195–212. Fresno: California State University.

Nemati, Fatemeh. 2010. Incorporation and complex predication in Persian. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG2010 Conference*, 395–415. Stanford: CSLI Publications.

Nunberg, Geoffrey, Ivan A. Sag & Thomas Wasow. 1994. Idioms. *Language* 70(3). 491–538.

Paul, Daniel. 2008. The individuating function of the Persian 'indefinite suffix'. In Simin Karimi, Vida Samiian & Donald Stilo (eds.), *Aspects of Iranian linguistics*, 309–328. Cambridge: Cambridge Scholars Press.

Samvelian, Pollet. 2012. *Grammaire des prédicats complexes – les constructions nom-verbe*. Paris: Lavoisier.

Samvelian, Pollet. 2018. Specific features of Persian syntax. In Anousha Sedighi & Pouneh Shabani-Jadidi (eds.), *The Oxford Handbook of Persian Linguistics*, 226–269. Oxford: Oxford University Press.

Samvelian, Pollet & Pegah Faghiri. 2014. Persian Complex Predicates: How Compositional Are They? *Semantics-Syntax Interface* 1(1). 43–74.

Samvelian, Pollet & Pegah Faghiri. 2016. Rethinking compositionality in Persian Complex Predicates. In *Proceedings of the Berkeley Linguistic Society 39th Annual Meeting, February 16-17 2013*, 212–227. Berkeley: Berkeley Linguistics Society.

Vahedi-Langrudi, Mohammad-Mehdi. 1996. *The syntax, semantics and argument structure of complex predicates in modern Farsi*. Ottawa: University of Ottawa dissertation.