



## Bilingualism, language contact & grammatical convergence: The case of complex predicates in Khuzestani Arabic

Seyyed Hatam Tamimi Sa'd\*

**Abstract.** A Complex Predicate (CP), e.g., ‘take a chance’, is a multi-word verbal construction that consists of a light verb (LV) and a nonverbal element (NVE) and that functions as a single semantic unit—i.e., together, both elements behave like a simplex verb. CPs are not historically productive in Khuzestani Arabic (KhA). Most KhA speakers are bilingual, speaking Persian as an L2, creating fertile ground for language contact phenomena. I provide evidence that KhA-Persian bilingual speakers increasingly form CPs by inserting native KhA light verbs *s‘ar* ‘become’ and *sæwwæh* ‘do / make’ into Persian complex-predicate templates and combining them with KhA items (e.g., adjectives or nouns), thereby creating innovative grammatical structures in response to asymmetrical language contact.

**Keywords.** Complex predicates, grammatical convergence, language contact, Khuzestani Arabic, Persian

**1. INTRODUCTION.** A Complex Predicate<sup>1</sup> (CP) is a construction that is composed of a Light Verb (LV) and a Nonverbal Element (NVE) or preverbal component. LVs, a term coined by Jespersen (1940), are, at least partially, semantically bleached verbs that lack thematic force in CP constructions and thus cannot function as predicates on their own. As a result, LVs need to combine with NVEs to form multiword expressions known as CPs.

- (1) a. *take a shower / a sneak / a plunge*  
b. *give a sigh / a shout / a ring*  
c. *have a cry / a read / a rest* (Butt 2010: 48)

CPs are thus idiomatic in that their meaning cannot be reconstructed from the meanings of the individual components; rather, their meanings must be learned by speakers (Faghiri & Samvelian 2021). CPs are productive in many languages such as Romance, Urdu, Persian, Kurdish, Bantu, Japanese, and Korean, among others. In CPs, the verbs still respect the verb + complement structure, for instance, in English. However, the idea is that light verbs are at least partially semantically devoid of their predicating force. For example, in (1a), *take a shower*, one does not actually physically *carry* a shower<sup>2</sup>. Rather, *take a shower* means ‘to shower’, with *take* in this construction serving as a verbal licenser (Butt 2010).

LVs are not, however, completely empty of any semantic content. Rather, they contribute some meaning. After all, *take a shower* and *give a shower* (e.g., *give the baby a shower*) have two different meanings. The term *light* means that the verb has a meaning that is hard to identify

---

\* I would like to thank the audience at the annual meeting of the Linguistic Society of America (LSA) 2026 for their valuable feedback and comments.

Author: Seyyed Hatam Tamimi Sa'd, Harry S Truman College ([shtamimi90@gmail.com](mailto:shtamimi90@gmail.com)).

<sup>1</sup> Also known as Compound Verbs, Composite Predicate, Light Verb Constructions (LVCs), or Serial Verb Constructions (SVCs) in the literature although there is no consensus on the appropriate term.

<sup>2</sup> Not in the literal sense, although one could imagine a situation in which a person physically carries a shower, for instance, at a hardware store.



d æz dæst dadæn  
 from hand give  
 ‘to lose’

To form CPs, Persian has several LVs including *fodæn* ‘to become’, *kærdæn* ‘to make / do’, *xordæn* ‘to collide / eat’, *dadæn* ‘to give’, *bordæn* ‘to take’, *kefidæn* ‘to pull’, *gereftæn* ‘to take’, *zædæn* ‘to hit’, among others, as shown in (5a-b). LVs, however, constitute a closed class while CPs are an open class (Korn 2013).

- (5) a. rændʒ kefidæn  
 suffering pull  
 ‘to suffer’
- b. zæmin xordæn  
 ground hit  
 ‘to fall’

This article explores the emergence of CPs in Khuzestani Arabic (KhA), a dialect of Arabic spoken in southwestern Iran, with a focus on their status and influence from Persian. I claim that, as a result of constant contact with Persian, KhA has used this language as a model to form CPs by using the two LVs, *sʿar* ‘he became’ and *sæwwæh* ‘he made / did’. Specifically, I present evidence that the emergence of CPs in KhA can be attributed to KhA speakers inserting native KhA light verbs (e.g., ‘do’ and ‘become’) into Persian complex-predicate templates and combining them with KhA items (e.g., adjectives or nouns).

1.1. KHUZESTANI ARABIC. Khuzestani Arabic (KhA) is a *gəlat* dialect (South Mesopotamian Arabic) spoken in the province of Khuzestan in southwestern Iran (Figure 1). It is a minority language spoken by Khuzestani Arabs. KhA is spoken alongside Persian (Farsi) in Khuzestan with almost all native speakers of KhA speaking Persian as a second language (L2). Despite its unique features, KhA has received little attention in the linguistic community due to limited academic access and sociopolitical marginalization (Matras & Shabibi 2007; Leitner 2022).



Figure 1. Khuzestan, Iran (Britannica)

Persian is the official language of Iran and is dominant in all formal domains, e.g., education, administration, and media, while KhA is a minority language spoken by between 2 and 3 million speakers<sup>5</sup> as their native language (L1) (Matras & Shabibi 2007). Thus, KhA carries far less prestige than Persian in formal domains and more KhA speakers feel some pressure to code-switch to Persian in everyday interactions, especially in formal and academic settings, formal

<sup>5</sup> The exact number of native speakers of KhA is difficult to obtain and verify for various reasons.

places such as hospitals and offices. KhA has been in intense contact with Persian and local Iranian minority languages (e.g., Luri, Bakhtiari, Qashqai Turkic, among others) for centuries (Figure 2). This asymmetrical bilingualism creates a setting conducive to contact-induced grammatical change.

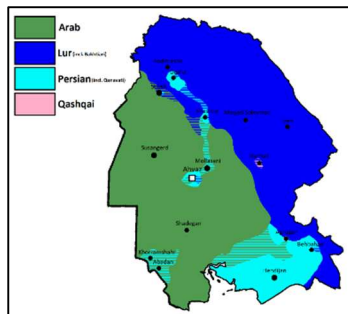


Figure 2. Geographic distribution of KhA and other languages in Khuzestan (Wikipedia)

Many KhA speakers live in Ahvaz (Ahwaz), the capital city of Khuzestan province (Figure 3). Although the intergenerational transmission within each ethnicity almost always occurs in that ethnic group’s respective language (e.g., KhA, Bakhtiari, etc.), the dominant language spoken in Ahvaz is Persian. Further, the population of Ahvaz has increased significantly over the last few decades due to a thriving natural gas and oil industry and the existence of important rivers such as Karun and Karkheh. Ahvaz is positioned near the border with Iraq and is home to a diverse population of 1.3 million people of various ethnicities including Persians, Arabs, Qashqai, Kurds, a dwindling Mandaic-speaking community, among others.



Figure 3. Ahwaz, capital city of Khuzestan (Wikipedia)

*Encyclopedia Iranica* dates the presence of Arabs in Khuzestan back to centuries before the Common Era. Following the advent of Islam, Arabic served as not only the language of everyday communication but also of literature, religion, scholarship, and administration in Khuzestan. This situation changed later in 1926 when Reza Shah Pahlavi came to power who favored Persian over other languages and who displaced many Arabs and resettled many Persians in Khuzestan (Matras & Shabibi 2007). Most Khuzestani Arabs today follow the Shia branch of Islam. Other religious minority groups in Khuzestan include Sunni Muslims, Christians, Jews, and Mandaeans. Like other ethnic groups, Khuzestani Arabs long taken pride in maintaining their customs and traditions—such as weddings, funerals, and dispute resolutions—and conducting these events largely in Arabic and wearing distinctive traditional attire for such occasions (Figure 4).



Figure 4. Khuzestani Arabs' traditional dance (Mahmoudian 2024)

**2. OBJECTIVES.** The current study examines CPs in KhA by investigating the role of contact with Persian in shaping KhA CPs. First, I provide descriptive information about KhA CPs. Next, I provide evidence that KhA productively uses Persian LVs to form genuine CPs. In particular, I examine two of the most productive LVs, the 'BECOME' type *sʕar* 'he became' and the 'DO' type *sæwwæh* 'he made / did'. I argue that these two verbs are modeled after their Persian counterparts, namely, *fodæn* and *kærdæn*, respectively, also two of the most productive LVs in Persian. In fact, Korn (2013) argues that the emergence of CPs in Persian is linked to the development of these two verbs.

**3. RESEARCH QUESTIONS.** The current study attempts to answer the following research questions:

- (i) What are the syntactic and semantic properties of CPs in KhA?
- (ii) What is the evidence that the combination of KhA LVs and NVEs forms genuine CPs?

**4. FINDINGS.** In this section, I aim to discuss two issues. First, I present some facts about CPs in KhA, focusing on their syntactic and semantic features. Next, I present evidence that the combination of the Persian LVs and KhA NVEs in KhA form genuine CPs. However, first I will briefly elaborate on *sʕar* 'he became' and *sæwwæh* 'he made / did' when used as heavy or thematic verbs in KhA as in (6) and (7).

(6) *sʕar-æt*                      *ʕæsʕæbi*  
 become-3SF.PRF angry  
 'She became angry.'

(7) *sæwwe-næh*      *zad*  
 make-1PL.PRF food  
 'We made food.'

4.1. SOME PROPERTIES OF CPS, NVES, & LVS IN KHA. In this section, I present some descriptive observations about CPs and their components, NVEs and LVs, to provide the reader with some preliminary insights prior to dealing with KhA CPs. This part addresses the first research question regarding the syntactic and semantic properties of CPs in KhA.

First, In KhA, the NVE in CPs may be modified (8) or scrambled (9). This suggests that the NVE may be syntactically independent of the LV:

(8) *ʕæli sæwwæh*              *tesʕadəf xætʕræh*  
 Ali make.3SM.PRF accident dangerous  
 'Ali had a dangerous accident.'

- (9) [tes<sup>ʕ</sup>adəf xæt<sup>r</sup>ræh]<sub>t<sub>i</sub></sub> ʕæli sæwwæh t<sub>i</sub>  
 accident dangerous Ali make.3SM.PRF

Next, both LVs *s<sup>ʕ</sup>ar* and *sæwwæh* host verbal features, e.g., tense, agreement, and negation (10a-c) and are fully conjugated according to KhA grammatical structure in terms of gender, number, and person. Thus, the LVs contribute verbal morphology and syntactic support as they host syntactic and semantic features and assign thematic roles while the NVEs contribute the lexical semantic content of the CP. In (10a), for instance, *t-sæwwi* introduces the agent / causer while *tes<sup>ʕ</sup>adəf* contributes the lexical concept ‘accident’.

- (10) a. t-sæwwi           tes<sup>ʕ</sup>adəf  
 FUT.3SF-make accident  
 ‘She will have an accident.’
- b. ma-t-sæwwi       tes<sup>ʕ</sup>adəf  
 NEG-FUT.3SF-make accident  
 ‘She will not have an accident.’
- c. \*t-sæwwi       ma-tes<sup>ʕ</sup>adəf  
 FUT.3SF-make NEG-accident  
 ‘She will not have an accident.’

Next, KhA frequently borrows lexical items, especially nouns, adjectives, and adverbs, from Persian. Thus, these LVs also frequently combine with nominals (11) or adjectivals (12) of Persian origin, forming ‘hybrid CPs’. This phenomenon especially holds true of speakers with stronger proficiency in Persian and weaker proficiency in KhA. In such CPs, loan words might be non-Persian, e.g., borrowed from English but through Persian. KhA hybrid CPs still follow the same template used in Persian. For instance, if Persian uses *kærdæn* with the non-Persian loan words, KhA uses *sæwwæh* as in (13).

- (11) sæww-æt       wærzæf  
 make-3SF.PRF exercise  
 ‘She exercised.’
- (12) s<sup>ʕ</sup>ar-æt       naraḥæh  
 become-3SF.PRF upset  
 ‘She became upset.’
- (13) sæww-æt       kəlik  
 make-3SF.PRF click  
 ‘She clicked.’

In addition, *sæwwæh* signals agentivity / cause while *s<sup>ʕ</sup>ar* denotes unaccusativity and result. Thus, *sæwwæh* takes an agentive subject (14) while the subject of *s<sup>ʕ</sup>ar* is a theme / patient (15 & 16). This shows that, in KhA as in Persian (Megerdooomian 2012), it is the verbal elements in CPs that contribute eventive properties, e.g., causation.

- (14) Kulumb   sæwwæh       ʔəmrikæh kæff  
 Columbus make.3SM.PRF America discovery  
 ‘Columbus discovered America.’

- (15) *sʕar*                      *tesʕadəf*  
 become.3SF.PRF accident  
 ‘There was / occurred an accident.’
- (16) *ʔəmrikæh sʕaræt*                      *kæjf*  
 America become-3SF.PRF discovery  
 ‘America was discovered.’

Finally, KhA NVEs are argument-bearing items consistent with Persian CPs (Karimi-Doostan 2005). Both *sæwwæh* and *sʕar* preserve their semantics and argument structure in KhA, in that *sæwwæh* denotes causative events (17) and takes one to three arguments (i.e., transitive and ditransitive).

- (17) *ma-y-sæwwi*              *tesʕadəf* (\**ma-tesʕadəf*)  
 NEG-FUT-make accident  
 ‘He will not have an accident.’

On the other hand, *sʕar* indicates involuntary eventive transitions and change-of-state (14) and typically takes a single argument (i.e., intransitive) (18):

- (18) *sʕar-æt*                      *bardæh*  
 become-3SF.PRF cold  
 ‘It became cold.’

Notably, *sʕar* allows KhA speakers to form passive-like structures (19 and 20). These passive structures at times mirror the exact opposite of the active voice counterparts (21 & 2).

- (19) *sʕar*                              *ʃæhid*  
 become.3SM.PRF martyr  
 ‘He was martyred.’
- (20) *ʔəmrikæh sʕar-æt*                      *kæjf*  
 America become-3SF.PRF discovery  
 ‘America became discovered (= was discovered).’
- (21) *sæww-et*              *əl-əktab tærdʒumæh*  
 make-1S.PRF the-book translation  
 ‘I translated the book.’
- (22) *əl-əktab sʕar*                              *tærdʒumæh*  
 the-book become.3SM.PRF translation  
 ‘The book was translated.’

This usage of *sʕar* allows KhA speakers to form anti-causative structures in an innovative way. KhA typically expresses passive or anti-causative meanings through a templatic derivation based on the Form VII pattern *infæʕæl*. For instance, the verb *kəʕæf* ‘he discovered’ corresponds to *inkəʕæf* ‘was discovered’, which follows the morphological template *infæʕæl* derived from the canonical root-and-pattern model *fæʕæl-æ*. This observation, i.e., use of *sʕar* to create passive-like structures also holds true in Persian (23). Such, seems to have been borrowed from, or at least influenced, by Persian:

- (23) *amrika kæff fod*  
 America discovery become.3S.PRF  
 ‘America was discovered.’

4.2. KHA CPS ARE GENUINE CPS. In this section, I present the following four lines of evidence to support that the combination of LVs borrowed from Persian and KhA elements form genuine CPs in KhA. This evidence is as follows.

First, KhA CPs may be replaced by simplex verbs (24), confirming their status as semantically unified predicates:

- (24) *sæwwa-ni næsʰihæh ↔ næsʰæh-ni*  
 make.3SM.PRF-me advice ↔ advise.3SM.PRF-me  
 ‘He gave me some advice / advised me.’

In (24), the CP *sæwwa-ni næsʰihæh* ‘gave me some advice’ may be replaced by the simple verb *næsʰæh-ni* ‘advised me’. The same may be done in Persian where a simple verb may replace a CP (25a-b). Furthermore, the verb *sæwwæh* is semantically bleached or thematically incomplete (Grimshaw & Mester 1988) since to advise does not literally mean ‘to make advice’. In this regard, then, CPs are said to be semantically noncompositional because the meaning of a CP is not predictable from its components (Goldberg 1996).

- (25) a. *kushesh kærdæn ↔ b. kushidæn*  
 attempt do  
 ‘to try’

(Examples 1a & 1b in Shabani-Jadidi 2014: 88)

The possibility of replacing a CP with a simple verb is also sociolinguistically interesting. In Persian, native speakers might find the simple verb archaic, poetic, literary, or formal (Megerdooian 2012; Shabani-Jadidi 2014). This is not necessarily the case in the context of KhA-Persian contact situation. Instead, KhA native speakers might find certain simple verbs such as *næsʰæh* ‘he advised’ more ‘accurate’, ‘original’, or ‘authentic’ than their CP counterparts and might view the use of their CP counterparts as ‘incorrect’ or indicative of one’s low proficiency in KhA. However, note that not all KhA CPs have a corresponding simple (i.e., heavy or thematic) verb.

Second, the same meaning can be conveyed in KhA by using a different verb that does not seem to behave similarly to a LV. For instance, in this case, *sæwwæh* may be replaced by *ætʰʰæ* ‘he gave’ (26).

- (26) *ætʰʰæ-ni næsʰihæh*  
 give.3SM.PRF-me advice  
 ‘He gave me some advice.’

Notably, the NVE *næsʰihæh* may not be specific for number in (27), i.e., when combined with *sæwwæh*, while it can be specific for number in (28), when used as an argument of *ætʰʰæ*. Instead, in (24), *næsʰihæh* is number neutral, i.e., neither singular nor plural.

- (27) ?? *sæwwa-ni næsʰihæht-iæn*  
 make.3SM.PRF-me advice-dual  
 ‘He gave me two pieces of advice.’

- (28)  $\text{æ}^{\text{t}}\text{t}^{\text{a}}\text{-ni}$                        $\text{næs}^{\text{c}}\text{i}^{\text{h}}\text{æht-iæn}$   
 give.3SM.PRF-me advice-dual  
 ‘He gave me two pieces of advice.’

In this sense,  $\text{næs}^{\text{c}}\text{i}^{\text{h}}\text{æh}$  in (24) does not pick up a reference in discourse. Further evidence comes from the fact that in (29),  $\text{næs}^{\text{c}}\text{i}^{\text{h}}\text{æh}$  cannot be co-indexed with a pronoun while in (30), it is more acceptable for  $\text{næs}^{\text{c}}\text{i}^{\text{h}}\text{æh}$  to be co-referenced with a pronoun.

- (29) \*  $\text{sæwwa-ni}$                        $\text{næs}^{\text{c}}\text{i}^{\text{h}}\text{æh}_i$   $\text{bəs}$   $\text{anæh}$   $\text{mɑ-xæði-æt-hæh}_i$   
 make.3SM.PRF-me advice      but I      NEG-take-1S-PRF-it  
 ‘He gave me some advice but I didn’t take it.’

- (30)  $\text{æ}^{\text{t}}\text{t}^{\text{æ}}\text{-ni}$                        $\text{næs}^{\text{c}}\text{i}^{\text{h}}\text{æh}_i$   $\text{bəs}$   $\text{anæh}$   $\text{mɑ-xæði-æt-hæh}_i$   
 give.3SM.PRF-me advice      but I      NEG-take-1S-PRF-it  
 ‘He gave me one piece of advice but I didn’t take it.’

The same prediction is borne out; that is, when  $\text{sæwwæh}$  is used as a heavy or thematic verb, it may be co-indexed with a pronoun (31) but not when it is used in a CP construction (24).

- (31)  $\text{sæwwe-næh}$        $\text{zad}_i$   $\text{bəs}$   $\text{mɑ-kælen-ah}_i$   
 make-1PL.PRF food but NEG-eat.3PL.PRF-it

I take the above findings as further evidence that  $\text{sæwwæh}$  behaves as an LV in (24) and that  $\text{æ}^{\text{t}}\text{t}^{\text{æ}}$  is a heavy or thematic verb at in (31). This suggests that  $\text{næs}^{\text{c}}\text{i}^{\text{h}}\text{æh}$  has different semantic and syntactic properties in CPs and in non-CP constructions, warranting different analyses. Specifically,  $\text{næs}^{\text{c}}\text{i}^{\text{h}}\text{æh}$  seems to behave like the internal argument of the thematic verb  $\text{æ}^{\text{t}}\text{t}^{\text{æ}}$  in (31) and as the NVE in (24). In any case, analyses of these elements have disagreed as whether preverbal nouns are internal arguments are different from non-specific objects and thus require different analyses (Mohammad & Karimi 1992) or whether they can be uniformly analyzed (Samvelian 2001).

Third, while Persian is a Subject-Object-Verb language (31a), inserting Persian-modeled LVs into KhA utterances does not change the canonical Subject-Verb-Object word order of KhA (31b). In Persian, CPs typically consist of an NVE followed by an LV, with the verb appearing in clause-final position. By contrast, when KhA expresses the same event using its native verbal resources, the verb precedes the object, rather than adopting the Persian SOV pattern, reflecting the language’s SVO order. This pattern aligns with cross-linguistic observations that syntactic structures are more resistant to borrowing than lexical items, particularly in the case of long-term contact.

- |      |   |         |
|------|---|---------|
| (31) | a. $\text{ʕæli}$ $\text{tæsədof}$ $\text{kærd}$<br>Ali   accident do.3S.PRF<br>‘Ali had an accident.’ | Persian |
|      | b. $\text{ʕæli}$ $\text{sæwwæh}$ $\text{tes}^{\text{c}}\text{ədof}$<br>Ali   make.3SM.PRF accident    | KhA     |

Finally, in KhA, just as in Persian, no constituent, e.g., adverbs or prepositional phrases, may intervene between the LV and the NVE. This restriction indicates that the LV and NVE form a single syntactically tight verbal complex. The ungrammaticality of (32) and (33) illustrates that inserting material between the LV and the NVE disrupts the CP and that they must

remain adjacent. This restriction on separability also exists in Persian CPs<sup>6</sup>. Obligatory adjacency between the LV and the NVE has been used as a diagnostic for CP formation (e.g., Nemati, Butt & King 2010; Karimi-Doostan 2011; Megerdooian 2012; Samvelian & Faghiri 2013). In fact, this inseparability is often taken as what distinguishes CPs from ordinary verb-complement syntactic combinations (Goldberg 1996).

(32) \* *mā-y-sæwwi l-yom tesʻadəf*  
 NEG-FUT-make today accident  
 ‘He will not have an accident today.’

(33) \* *ʔəmrikæh sʻar-æt əb-sənæt 1492 kæff*  
 America become-3SF.PRF in-year 1492 discovery  
 ‘America was discovered.’

**5. DISCUSSION & CONCLUSION.** KhA and Persian have co-existed for centuries in almost every setting in Khuzestan, Iran. However, Persian is the dominant language, carrying more prestige due to its status as the official language while KhA is non-dominant due to its status as a minority language. I investigated whether KhA has genuine CPs and whether the existence of CPs is due to contact with Persian. I presented four lines of evidence that KhA has true CPs. First, a CP in KhA may be replaced by a simple, thematic verb (if one is available). Second, the same meaning is conveyed by the simplex verb as by the CP. Third, the use of the two LVs *sæwwæh* and *sʻar* in KhA does not change the word order of the KhA string. Finally, the LV and NVE form a contiguous unit that do not allow separability.

Findings showed that KhA seems to have modeled its CPs on the semantic structure of Persian CPs, a language that uses CPs extensively. The emergence of CPs in KhA can be analyzed as pattern replication (Matras & Sakel 2007), whereby bilingual speakers adopt the abstract structural template of a construction from one language while filling it with lexical material from another. In this case, KhA speakers replicate the Persian CP structure [NVE + LV] while inserting native KhA LVs and lexical elements.

The fact that the two LVs *sæwwæh* and *sʻar* form a large number of CPs in KhA is also worth investigating. In this regard, drawing on Korn (2013), I argue that it is possible that these two verbs have triggered the emergence of CPs in KhA thanks to their features as being ‘underspecified’, i.e., semantically bleached. According to Korn, the two Persian verbs *kærdæn* ‘do’ and *fodæn* ‘become’ have the features Instigation [+INST] and Affectedness [+AFF], respectively, making them characteristic for prototypical agents and patients. These features thus explain the transitivity of *sæwwæh* and intransitivity / passivity of *sʻar*.

The current study is the first attempt to understand contact-induced complex predicate formation in KhA and the role of contact with Persian. Future research should investigate the factors that influence the choice of the use of CPs (as opposed to simple verbs) in KhA, including the speakers’ age, proficiency in both KhA and Persian, level of education, place of residence (e.g., Persian-dominated vs. KhA-dominated), attitudes towards each language, among others. Despite their increasing popularity, using CPs is not always welcomed in KhA. Instead, to some speakers, Persian-modeled CPs may be perceived as signaling lower proficiency or as one attempting to sound more Persian than Arab. In this regard, some might describe someone’s KhA as Persianized.

---

<sup>6</sup> For some exceptions, see Ghomeshi (1996), Samvelian (2012, 2018), and Faghiri and Samvelian (2021).

The syntactic and semantic properties of both LVs and NVEs are also another area that warrant extensive research. Furthermore, I anticipate that other LVs in KhA also behave similarly to the ones investigated here, namely, *s'ar* and *sæwwæh*, and are semantically modeled according to Persian CP structure. Yet, further research is needed to confirm this prediction. The following is an example of a different LV in KhA which the CP semantics seems to have been directly borrowed from Persian.

- (34) bætʃfæ-ro sili zæd Persian  
 boy-OBJ slap hit.3S.PRF  
 ‘He/She slapped the boy.’
- (35) tʰægg əl-wələd əsʰtʰar<sup>7</sup> ↔ sʰətʰær əl-wələd KhA  
 hit.3SM.PRF the-boy slap ↔ slap.3SM.PRF the-boy  
 ‘He slapped the boy.’

The emergence of CPs in KhA illustrates how intense bilingualism can result in the replication of abstract grammatical constructions across unrelated languages, in this case, Persian and Arabic, two typologically different languages. KhA speakers adopt the Persian CP template without adopting Persian word order, indicating that structural borrowing may target semantic and morphosyntactic constructions independently of clausal level syntax.

Similar processes of LV borrowing have been documented in other bilingual contexts, e.g., Turkish / Dutch (Backus 1992) and Spanish / German (González-Vilbazo & López 2011), though framed as code-switching. However, since in our case, the entire clause remains in KhA, the KhA-Persian case involves structural borrowing, not code-switching. This case exemplifies structural borrowing under long-term bilingualism, contributing to broader discussions of contact-induced grammatical convergence (Matras & Sakel 2007).

Future research needs to address the contact between Persian and other languages spoken in Iran and whether this contact has resulted in the emergence of CPs as it has in KhA. It is also possible that CPs in KhA are semi-productive light verb constructions as they have just started to emerge. Further comparisons between KhA and Persian can be telling as CPs are fully grammaticalized and are highly productive in Persian.

## References

- Backus, Ad. 1992. *Patterns of language mixing: A study of Turkish-Dutch bilingualism*. Wiesbaden: Harrassowitz.
- Butt, Miriam. 2010. The light verb jungle: Still hacking away. In Mengistu Amberber, Brett Baker & Mark Harvey (eds.), *Complex predicates: Cross-linguistic perspectives on event structure*, 48–78. Cambridge: Cambridge University Press.
- Dabir-Moghaddam, Mohammad. 1997. Compound verbs in Persian. *Studies in the Linguistic Sciences* 27(2). 25–59.
- Encyclopaedia Britannica. Khuzestan. 2026. <https://www.britannica.com/place/Khuzestan>.
- Faghiri, Pegah & Pollet Samvelian. 2021. The issue of “separability” in Persian complex predicates. In Berthold Crysmann & Manfred Sailer (eds.), *One-to-many relations in morphology, syntax, and semantics*, 117–149. Berlin: Language Science Press. DOI: 10.5281/zenodo.4729799

<sup>7</sup> Note, however, that the NVE əsʰtʰar ‘slap’ in tʰægg əl-wələd əsʰtʰar maybe numbered: əsʰtʰar-ian ‘two slaps / twice’. This exception warrants a separate investigation.

- Ghomeshi, Jila. 1996. *Projection and inflection: A study of Persian phrase structure*. Ontario, Canada: University of Toronto dissertation.
- Goldberg, Adele E. 1996. Words by default: Optimizing constraints and the Persian complex predicate. In *Proceedings of the 22<sup>nd</sup> Annual Meeting of Berkeley Linguistic Society*, 132–146. Berkeley: Linguistic Society of America.
- González-Vilbazo, Kay & Luis López. 2011. Some properties of light verbs in code-switching. *Lingua* 121. 832–850.
- Grimshaw, Jane & Armin Mester. 1988. Light verbs and theta-marking. *Linguistic Inquiry* 19(2). 205–232.
- Jespersen, Otto. 1940. *A modern English grammar on historical principles*. London: Allen & Unwin.
- Karimi-Doostan, Gholamhosein. 1997. *Light verb constructions in Persian*. Colchester, England: University of Essex dissertation.
- Karimi-Doostan, Gholamhosein. 2011. Separability of light verb constructions in Persian. *Studia Linguistica* 65. 70–95. <https://doi.org/10.1111/j.1467-9582.2011.01178.x>
- Korn, Agnes. 2013. Looking for the middle way: Voice and transitivity in complex predicates in Iranian. *Lingua* 135. 30–55.
- Leitner, Bettina. 2022. *Grammar of Khuzestani Arabic*. Leiden: Brill.
- Mahmoudian, Maryam. 2024. Iran ethnic groups: Beyond Persians – A guide to diversity. Manuscript. <https://www.eavartravel.com/blog/2024/5/11/160868/iran-ethnic-groups/>.
- Matras, Yaron & Maryam Shabibi. 2007. Grammatical borrowing in Khuzistani Arabic. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 137–149. Berlin: Mouton de Gruyter.
- Matras, Yaron & Jeanette Sakel. 2007. Introduction. In Yaron Matras & Jeanette Sakel (eds.), *Grammatical borrowing in cross-linguistic perspective*, 1–13. Berlin: Mouton de Gruyter.
- Megerdooomian, Karine. 2001. Event structure and complex predicates in Persian. *Canadian Journal of Linguistics* 46. 97–125.
- Mohammad, Jan & Simin Karimi. 1992. Light verbs are taking over: Complex verbs in Persian. *Proceedings of WECOL* 5. 195–212. Fresno, CA: California State University.
- Natel Khanlari, Parviz. 1973. *Tarix-e zaban-e Farsi (The history of the Persian language)*, Vol. 2. Tehran: Bonyad-e Farhang.
- Nemati, Fatemeh, Miriam Butt & Tracy Holloway King. 2010. Incorporation and complex predication in Persian. *Proceedings of LFG10*. 374–394.
- Samvelian, Pollet. 2001. Le statut syntaxique des objets nus en persan. *Bulletin de la Société de Linguistique de Paris* XCVI. 349–388.
- Samvelian, Pollet. 2018. Specific features of Persian syntax: The Ezafe construction, differential object marking, and complex predicates. In Anousha Sedighi & Pouneh Shabani-Jadidi (eds.), *The Oxford Handbook of Persian Linguistics*, 226–270. Oxford: Oxford University Press.
- Samvelian, Pollet & Pegah Faghiri. 2013. Introducing PersPred: A syntactic and semantic database for Persian complex predicates. *Proceedings of the 9<sup>th</sup> Workshop on Multiword Expressions*, 1120. Association for Computational Linguistics.
- Shabani-Jadidi, Pouneh. 2014. *Processing compound verbs in Persian: A psycholinguistic approach to complex predicates*. Netherlands: Amsterdam University Press. <https://doi.org/10.1515/9789400601758>.

Soltani, R., Mohammad Amouzadeh & Hadaegh Rezaei. 2024. The Persian light verb *dādan* ‘to give’: Causation and more. *Journal of Researches in Linguistics* 16(2). 1–18. DOI: 10.22108/jrl.2024.140019.1822

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

Wikipedia. Ahvaz. 2026. <https://en.wikipedia.org/wiki/Ahvaz>