



HAL
open science

Colloquial Persian: Towards a New Rise of Simple Verbs?

Dorian Pastor, Agnes Korn, Christian Rammer

► **To cite this version:**

Dorian Pastor, Agnes Korn, Christian Rammer. Colloquial Persian: Towards a New Rise of Simple Verbs?. *Transactions of the Philological Society*, 2024, 122 (1), pp.21-48. 10.1111/1467-968X.12270 . hal-04128873

HAL Id: hal-04128873

<https://hal.science/hal-04128873>

Submitted on 17 Mar 2024

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.

Colloquial Persian: Towards a New Rise of Simple Verbs?¹

By DORIAN PASTOR,¹ AGNES KORN² & CHRISTIAN RAMMER³
¹École Normale Supérieure, Paris, ²CNRS, Paris; ³ZEW, Mannheim

ABSTRACT

For the formation of new verbs, Persian very productively employs complex predicates, while the derivation of simple verbs has been held to be marginal or even inexistent. We use data from internet fora, blogs and messaging services to argue that the derivation of simple verbs from nominals is a productive process at least among speakers of colloquial Persian who actively use online media. Two acceptability tests of seven verbs and two ‘placebo verbs’ conducted among 58 participants produced more than 7,000 gradings. They show that such verbs are surprisingly well accepted, with statistically significant differences depending on age, gender, region, TAM form and the individual verbs. Some of these results seem quite straightforward (e.g., acceptance declining with age), others rather not (for instance, the present and past progressive is very well accepted). The results run counter to the assumption that a gradual substitution of synthetic structures (simple verbs) by analytical ones (complex predicates) is a unidirectional process.

La formation de nouveaux verbes en persan emploie de manière très productive les prédicats complexes, tandis que la dérivation de verbes simples a été jugée marginale, voire inexistante. Dans cet article, nous nous appuyons sur des données issues de l’Internet (forums, blogs et services de messagerie divers) pour affirmer que la création de verbes simples à partir de noms est un processus productif, tout du moins parmi les locuteurs du persan vernaculaire qui utilisent activement les réseaux sociaux. Deux tests d’acceptabilité de sept verbes et de deux « verbes placebos » menés auprès de 58 participants ont permis d’obtenir plus de 7000 notes. Ils montrent que ces verbes sont assez bien acceptés, avec des différences statistiquement significatives selon l’âge, le sexe, la région, la forme TAM et les verbes individuels. Certains des résultats semblent relativement attendus (par exemple, l’acceptation diminue avec l’âge) mais d’autres le sont moins (par exemple, le présent et le passé progressifs sont très bien acceptés). Les résultats vont à l’encontre de l’hypothèse selon laquelle une substitution progressive des structures synthétiques (les verbes simples) par des structures analytiques (les prédicats complexes) est un processus unidirectionnel.
(French)

1. COMPLEX VERBS AND SIMPLE VERBS IN PERSIAN

1.1. Introduction and state-of-the-art

As stated by Telegdi (1951) and numerous other linguists, Persian has a very restricted number of simple verbs, and many authors add that simple verbs are a closed class. Some works suggest an absolute figure of simple verbs, or provide an exhaustive list. Correspondingly, there is a very high number of complex predicates, which consist of a non-verbal element and a so-called light verb, for example, *harf zadan* ‘to talk (letter + hit)’, *javāb dādan* ‘to answer (answer +

¹ The idea for this article as well as the design of the questionnaires and the interviews with the respondents were all carried out by D.P. The statistical calculations and the figures are by Ch.R., the article text is largely by A.K.

give)', *dorost kardan* 'to prepare (right + do)'. In some cases, there is a complex predicate and a simple verb with the same meaning (e.g., *motevaje šodan / fahmidan* 'to understand'), but most complex predicates do not have a simple verb as equivalent (see Mir-Djalali 1974: 11). This renders complex predicates a phenomenon of another category than structurally parallel patterns in languages such as German or French (Table 1).

TABLE 1. The status of complex predicates vs. simple verbs

	Complex predicate	Simple verb
German	<i>Antwort geben</i>	<i>antworten</i>
French	<i>donner une réponse</i>	<i>répondre</i>
Persian	<i>javāb dādan</i>	∅

In a language having complex predicates, the pattern can become very productive; it permits to combine any nominal (including borrowed ones) with a semantically fitting light verb (Butt 2010: 52). This may contribute to simple verbs becoming more and more marginalised.

One might thus expect that any newly formed verb will be a complex predicate, and that new simple verbs will not be formed in a language such as Persian. Indeed, it has often been stated that complex predicates are 'the only productive process used in creating denominative verbs from Persian lexemes or loanwords from other languages' (Ciancaglini 2011: 3).²

Such a statement is perhaps somewhat too categorical, at least for Persian as a whole: as implied by Windfuhr & Perry (2009: 447), inherited or borrowed nouns have also been used as present stems (and a past stem is formed with the productive suffix *-id*, see below), even if 'this procedure is no longer very productive'; they cite, for example, *nāmidan* 'to call (by a name)' and *fahmidan* 'to understand' mentioned above, derived from *nām* 'name' and Arabic *fahm* 'understanding', respectively.³ However, it is certainly true that complex predicates are the common and productive way of forming new verbs in Persian.

Batani (1990: 69) notes that a couple of simple verbs have emerged in the last decades in the colloquial language, which are derived from substantives and adjectives, but he nevertheless concludes that simple verbs are not productive and not generally accepted. These new simple verbs often have a complex predicate equivalent (Table 2).

[p. 23]

TABLE 2. Some examples of new simple verbs in Persian (Batani 1990: 69)⁴

Simple verb	Meaning	Complex predicate
<i>lāsīdan</i>	'to flirt'	<i>lās zadan</i> (flirt hit)
<i>šalīdan</i>	'to limp'	<i>šal zadan</i> (limp hit)
<i>šutīdan</i>	'to shoot (football)' ⁵	<i>šut kardan</i> (shoot do)
<i>sorīdan</i>	'to slide'	<i>sor x^wordan</i> (slippery eat)
<i>tiġīdan</i>	'to defraud'	<i>tiġ zadan</i> (blade hit) ⁶

The usual practice of citing Persian verbs by the infinitive should not mislead readers into considering *-īdan* a suffix by which verbs are derived. In fact, the new simple verbs are formed by conversion of a noun into a present stem (i.e., the noun is identical to the present stem), for

² Similarly, Lazard (2006: 279) and many other authors.

³ For *nāmidan*, see also Section 5.4.1.

⁴ English translations and complex predicate equivalents are ours.

⁵ For *šutīdan*, see Section 2.3.

⁶ The complex predicate *tiġ zadan* usually means 'to shave'. (For Classical Persian, Hayyim (1933-1935) translates it 'to strike with a sword'.) In colloquial Persian, the verb has the metaphorical meaning of getting money from someone by tricks and lies, which is the only sense of *tiġīdan*.

example, *tiġ* ‘blade’ → *tiġ-* defraud.PRS, *mi-tiġ-am* ‘I defraud’.⁷ Just like any new or secondary verb, these verbs are assigned a past stem by suffixing the past stem morpheme *-id* (*tiġ-id* defraud-PST),⁸ from which in turn the infinitive is derived by adding *-an*.⁹ Table 3 shows the inflection of such a verb.

TABLE 3. Some basic morphology of Persian verbs (example: *gugelidan* ‘to google’)¹⁰

	Present		Past	
	Singular	Plural	Singular	Plural
1	<i>mi-gugel-am</i>	<i>mi-gugel-im</i>	<i>gugel-id-am</i>	<i>gugel-id-im</i>
2	<i>mi-gugel-i</i>	<i>mi-gugel-in</i>	<i>gugel-id-i</i>	<i>gugel-id-in</i>
3	<i>mi-gugel-e</i>	<i>mi-gugel-an</i>	<i>gugel-id-∅</i>	<i>gugel-id-an</i>
nonfinite forms	PRF <i>gugel-id-e</i>		INF <i>gugel-id-an</i>	

1.2. New data

Our own investigations¹¹ give at hand that additional new simple verbs are emerging in Persian, including examples such as (1) and (2).

[p. 24]

- (1) *ne-mi-x^wād* *bāhām* *be-ħarf-e*
 NEG-IPFV-want.PRS.3SG with.PC1SG SBJV-talk.PRS-3SG
 ‘S/he doesn’t want to talk with me.’ (WEB)
- (2) *man* *bāyad* *be-zang-am* *be* *mādar* *šōhar=am?*
 I must SBJV-call.PRS-1SG to mother husband=PC1SG
 ‘Do I have to call my husband’s mother?’ (WEB)

Examples (3–7) illustrate some additional TAM forms: present and past progressive (3, 4), subjunctive (5, see also 2, 11 and 14), past tense (4, 6, 12) and imperfect (7). An example for the imperative is found in (11).

- (3) *did-am* *dār-e* *mi-barf-e*
 see.PST-1SG have.PRS-3SG IPFV-snow.PRS-3SG
 ‘I saw that it was snowing.’ (WEB)

⁷ A few verbs pattern differently in being based on a noun in *-e* or *-i*, which is deleted in the conversion, see Section 2.3.

⁸ According to Korn (2009: 198f, 209), *-id* (Middle Persian *-īd*) derives from Proto-Iranian **-i-ta-*. In inherited verbs, the suffix is used to derive morphologically transparent past stems, replacing items that have become opaque, e.g., *ruy-id* ‘grown’ (replacing *rost*); furthermore, *-id* serves to recruit a past stem for verbs which do not have one, for example, *ras-id* ‘arrived’ or formations from loanwords such as *fahm-id* (Korn 2009: 198). Besides *-id*, there are also some past stems in *-d/t* and *-ād*.

⁹ Note that the present indicative and subjunctive are far better accepted than the past tense (see Section 4.6), which could imply that the present stem, rather than the past stem, is considered as basic in the mind of the speakers.

¹⁰ For the verbal categories as applied in this paper, see Table 8.

¹¹ The examples presented in this article are from written sources found on internet fora and blogs – some of them over 10 years old – (‘WEB’ in what follows) and messaging services (‘MES’: SMS, WhatsApp, Telegram). When speakers use colloquial Persian in writing (as in our examples), the orthography (using the Persian alphabet) is to some extent adjusted to colloquial pronunciation, but not in all cases. Most instances of standard Persian *ān* are pronounced *un* in colloquial style, which is usually written as such. Similarly, the ending of the 3rd singular is noted *-e* as it is pronounced (vs. *-ad* of the Standard language). Conversely, *ba’d-az* ‘after’ is written as in the Standard language (with ‘ein) although it is pronounced *ba:d* in colloquial style. – We note the pronominal clitics (PC) and the marker *=(r)o / =rā* (for which see note 12) attached to the preceding word by =, but the subordinator *ke*, the connector *o* ‘and’, *ham* ‘also’ and the forms of the copula as separate words.

- (4) *do sā'at e dāšt-am mi-tāypid-am,*
 two hour COP3SG have.PST-1SG IPFV-type.PST-1SG
yeho hama=š pākid
 suddenly all=PC3SG delete.PST(3SG)
 'I was typing for two hours. Suddenly it all got deleted.' (WEB)
- (5) *sa'y mi-kon-am bištar bāš-am*
 try IPFV-do.PRS-1SG more be.SBJV-1SG
ke del=etun na-tang-e
 SUB heart=PC2PL NEG-tight.SBJV-3SG
 'I try to be more (present) so that you don't miss [me].' (WEB)
- (6) *dam=et garm rafiq ke javābid-i*
 breath=PC2SG warm friend SUB answer.PST-2SG
 'Thanks for having answered, dude.' (WEB)
- (7) *kāški parastār-e ye kam bištar mi-šuxid*
 if_{only} nurse-REF one bit more IPFV-joke.PST(3SG)
 'If only the nurse had joked a bit more!' (WEB)

In some cases, speakers voice their impressions about the use of this or that simple verb. For instance, they say that *fekridan* 'to think', *gušidan* 'to listen', and *darsidan* 'to study' are more likely to be used by teenagers, and refer to typical situations they find themselves in while the other verbs are not restricted to a particular age group. This indicates that a sociolinguistic study of the simple verbs might be interesting.

In many cases, the simple verb corresponds to one complex predicate. However, sometimes this is not so straightforward, which raises interesting questions for the morphosyntactic side of these verbs. In the case of *gušidan* 'to listen', there are two complex predicates that likewise mean 'listen', viz. *guš kardan* (ear + do) and *guš dādan* (ear + give), of which the former takes a direct object (9) while the latter takes a complement with the preposition *be* (10). A speaker explained that she would not accept *gušidan* with *be* (8b), but would prefer the topic marker *=(r)o*¹² (8a); in this sense, *gušidan* thus corresponds to *guš kardan*.¹³

[p. 25]

- (8a) *āhang=o mi-guš-e*
 song=RĀ IPFV-listen.PRS-3SG
- (8b) ? *be āhang mi-guš-e*
 to song IPFV-listen.PRS-3SG
 'S/he listens to the song.'
- (9) *zamān=i ke ġamgin bāš-am in āhang=rā guš mi-kon-am*
 time=IDV SUB sad be.SBJV-1SG DEM1 song=RĀ ear IPFV-do.PRS-1SG
 'When I'm sad, I listen to this song.'
- (10) *tanhā be yek āhang guš dād-e*
 only to one song ear give.PST-PRF(3SG)
 'S/he listened to only one song.'

¹² The clitic *=rā* has various functions, among which the marking of definite direct objects is the most frequent. The clitic has been analysed as 'marking DPs that are both specific and assigned accusative case' (Karimi & Smith 2020: 169), but also as topic marker (e.g., Ghomeshi 1997). For the purposes of this paper, we will gloss it =RĀ.

¹³ It seems that *guš kardan* (ear + do) implies a greater attention and that the imperative *guš kon!* carries more authority and urgency than *guš bedeh!*

The rise of simple verbs follows a period in which complex predicates tend to be replaced by complex predicates, and the latter have widely been regarded as the only productive way of forming new verbs in Persian. The phenomenon is thus of interest also beyond the study of Persian since it runs counter to the overall tendency observable in Persian, and also cross-linguistically, by which synthetic constructions are replaced by analytical ones (see Section 5.4).

2. DISCOVERING SIMPLE VERBS

2.1. Inventory of new simple verbs

Drawing on social media, as well as conversations with friends and acquaintances, we collected the simple verbs listed in Table 4.

TABLE 4. Alphabetical list of new simple verbs observed¹⁴

	Simple verb	Complex predicate	
With one complex predicate equivalent			
<i>kardan</i> ‘to do’	<i>‘aṣīdan</i>	‘to sneeze’ ¹⁵	<i>‘aṣe kardan</i>
	<i>bağalīdan</i>	‘to hug’	<i>bağal kardan</i>
	<i>baḥṣīdan</i>	‘to have an argument’	<i>baḥṣ kardan</i>
	<i>bāzīdan</i>	‘to open’	<i>bāz kardan</i>
	<i>čatīdan</i>	‘to chat’ ¹⁵	<i>čat kardan</i>
	<i>fekrīdan</i>	‘to think’	<i>fekr kardan</i>
	<i>gugelīdan</i>	‘to google’ ¹⁵	<i>gugel kardan</i>
	<i>gušīdan</i>	‘to listen’	<i>guš kardan</i>
	<i>jišīdan</i>	‘to urinate’	<i>jiš kardan</i>
	<i>kašfīdan</i>	‘to discover’	<i>kašf kardan</i>
	<i>kilikīdan</i>	‘to click’ ¹⁵	<i>kilik kardan</i>
	<i>komakīdan</i>	‘to help’	<i>komak kardan</i>
	<i>kopīdan</i>	‘to copy’ ¹⁵	<i>kopi kardan</i>
	<i>kuftīdan</i>	‘to eat’ (colloquial)	<i>kuft kardan</i>
	<i>loxtīdan</i>	‘to undress (tr.)’	<i>loxt kardan</i>
	<i>našbīdan</i>	‘to install’	<i>našb kardan</i>
	<i>negāhīdan</i>	‘to look (at)’	<i>negāh kardan</i>
	<i>pārkiđan</i>	‘to park’ ¹⁵	<i>pārkiđan kardan</i>
	[p. 26] <i>pāsiđan</i>	‘to pass (an exam)’ ¹⁵	<i>pās kardan</i>
	<i>paxšiđan</i>	‘to play (music, film)’	<i>paxš kardan</i>
	<i>qahrīdan</i>	‘to sulk, to be mad’	<i>qahr kardan</i>
	<i>qomāriđan</i>	‘to gamble’	<i>qomār kardan</i>
	<i>šabriđan</i>	‘to wait’	<i>šabr kardan</i>
	<i>šoḥbatīdan</i>	‘to discuss’	<i>šoḥbat kardan</i>
	<i>serčiđan</i>	‘to search (on web)’ ¹⁵	<i>serč kardan</i>
	<i>šuxīdan</i>	‘to joke’ ¹⁶	<i>šuxi kardan</i>
	<i>ta‘ārofīdan</i>	‘to kindly offer’	<i>ta‘ārof kardan</i>
<i>talafīdan</i>	‘to waste’	<i>talaf kardan</i>	

¹⁴ These verbs have come to our attention since we started our study in mid-2019. The list is surely incomplete. Some additional items cited by Bateni 1990 are found in Table 2.

¹⁵ See Section 2.3 below.

¹⁶ See Sections 2.2 and 2.3 below.

	<i>tāypidan</i>	‘to type’ ¹⁵	<i>tāyp</i>	<i>kardan</i>
	<i>varzešidan</i>	‘to make sport’	<i>varzeš</i>	<i>kardan</i>
	<i>zoqidan</i>	‘to rejoice’	<i>zoq</i>	<i>kardan</i>
<i>zadan</i> ‘to hit’	<i>gapidan</i>	‘to chat’	<i>gap</i>	<i>zadan</i>
	<i>goridan</i>	‘to nag, to grunt’	<i>gor</i>	<i>zadan</i>
	<i>harfidan</i>	‘to talk’	<i>harf</i>	<i>zadan</i>
	<i>imeylidan</i>	‘to send an email’ ¹⁵	<i>imeyl</i>	<i>zadan</i>
	<i>qadamidan</i>	‘to walk’	<i>qadam</i>	<i>zadan</i>
	<i>zangidan</i>	‘to call (telephone)’	<i>zang</i>	<i>zadan</i>
<i>dādan</i> ‘to give’	<i>javābidan</i>	‘to answer’	<i>javāb</i>	<i>dādan</i>
	<i>leftidan</i>	‘to leave (an online group)’ ¹⁵	<i>left</i>	<i>dādan</i>
	<i>gāzidan</i>	‘to accelerate (a car)’ ¹⁵	<i>gāz</i>	<i>dādan</i>
	<i>tarjihidan</i>	‘to prefer’	<i>tarjih</i>	<i>dādan</i>
	<i>xabaridan</i>	‘to inform, to let know’	<i>xabar</i>	<i>dādan</i>
<i>xundan</i> ‘to read’	<i>darsidan</i>	‘to study’	<i>dars</i>	<i>xundan</i>
<i>kešidan</i> ‘to pull’	<i>derāzidan</i>	‘to lie down’	<i>derāz</i>	<i>kešidan</i>
<i>gereftan</i> ‘to take’	<i>aksidan</i>	‘to take a picture’	<i>aks</i>	<i>gereftan</i>
	<i>dušidan</i>	‘to take a shower’	<i>duš</i>	<i>gereftan</i>
	<i>jašnidan</i>	‘to celebrate’	<i>jašn</i>	<i>gereftan</i>
<i>xordan</i> ‘to eat/drink’	<i>čāyidan</i>	‘to have tea’	<i>čāy</i>	<i>xordan</i>
	<i>qahvidan</i>	‘to have coffee’ ¹⁷	<i>qahve</i>	<i>xordan</i>
<i>dāštan</i> ‘to have / hold’	<i>dustidan</i>	‘to love’	<i>dust</i>	<i>dāštan</i>
	<i>šakidan</i>	‘to doubt’	<i>šak</i>	<i>dāštan</i>
<i>šodan</i> ‘to become’	<i>pašimunidan</i>	‘to regret’	<i>pašimun</i>	<i>šodan</i>
	<i>tangidan</i>	‘to become tight’	<i>tang</i>	<i>šodan</i>
<i>umadan</i> ‘to come’ ¹⁸	<i>barfidan</i>	‘to snow’	<i>barf</i>	<i>umadan</i>
With two complex predicate equivalents				
<i>kardan</i> ‘to do’ / <i>šodan</i> ‘to become’	<i>āgāzidan</i>	‘to start’ (tr.)	<i>āgāz</i>	<i>kardan</i>
		‘to start’ (itr.)	<i>āgāz</i>	<i>šodan</i>
	<i>avazidan</i>	‘to change (tr.)’	<i>avaz</i>	<i>kardan</i>
		‘to be changed’	<i>avaz</i>	<i>šodan</i>
	<i>pākidan</i>	‘to delete (file, message)’	<i>pāk</i>	<i>kardan</i>
		‘to be deleted’	<i>pāk</i>	<i>šodan</i>
<i>qabulidan</i>	‘to accept’	<i>qabul</i>	<i>kardan</i>	
	‘to be accepted’	<i>qabul</i>	<i>šodan</i>	
<i>tamumidan</i>	‘to complete’	<i>tamum</i>	<i>kardan</i>	
	‘to be completed’	<i>tamum</i>	<i>šodan</i>	
<i>oftādan</i> ‘to fall’ / <i>raftan</i> ‘to go’	<i>rāhidan</i>	‘to leave’	<i>rāh</i>	<i>oftādan</i>
		‘to walk’	<i>rāh</i>	<i>raftan</i>

Broadly agreeing with the distribution of light verbs in complex predicates, by far the largest group of these verbs corresponds to complex predicates with *kardan* ‘to do’. To a lesser extent, they correspond to complex predicates with *zadan*, ‘to hit’. Other complex predicates [p. 27] with *dādan* ‘to give’, *xundan* ‘to read’, *gereftan* ‘to take’, *xordan* ‘to eat/drink’, *dāštan* ‘to have’ and *šodan* ‘to become’ are less common as corresponding to new simple verbs.

¹⁷ See Section 2.3. *qahvidan* also exists in the meaning ‘make/prepare coffee’ (*qahve tōlid kardan*).

¹⁸ The pattern is *barf miyād / umad* ‘snow comes / came’ for ‘it snows / snowed’.

2.2. New or not new?

Some of the verbs in Table 4 are found in Persian dictionaries, but are nevertheless perceived as being new formations by speakers.

The following items observed in our study show semantics different from those noted in the literature, or from the apparently corresponding complex predicate, suggesting that they are new formations:

- *āgāzidan* is noted in Hayyim (1933-1935) as the counterpart of *āgāz kardan* ‘to start (tr.)’, but not of *āgāz šodan* ‘to start (itr.)’.
- *pākidan* means ‘to clean, to purify’ in Dehxodā (*setordan; ruftan*), while in our data, *pākidan* is commonly used with the meaning ‘to delete / be deleted (e.g., a message on the Internet)’ (11).

- (11) *tā* *sā’at=e* *punzdah* *be-pāk-in,*
 until hour=EZ fifteen IMP-delete.PRS-2PL
na-pāk-in *mi-left-am.*
 NEG-delete.SBJV-2PL IPFV-leave.PRS-1SG
 ‘Delete until fifteen (you have until 3pm to delete [it]),
 (if) you don’t delete [it], I will leave (the online group).’ (WEB)

- *paxšidan* means ‘to become crushed; to spread, to widen’ according to Dehxodā (*kufte šodan, pahn gardidan, parākandan*); in our data, it is commonly used for ‘to play (switch on) a song or film (for oneself or others)’, see (12).

- (12) *šad* *hezār* *bār* *paxšid-am=eš*
 hundred thousand time transmit.PST-1SG=PC3SG
 ‘I played it (i.e. listened to the song) hundreds of thousands of times.’ (WEB)

- *goridan* is ‘to raise the voice; to scream; to roar (lion etc.), to grumble (thunder)’ in Dehxodā, while in our data *goridan* is used with the meaning ‘to nag, to grunt’.

- Steingass 1892 (dictionary of classical Persian) has *šōxīdan* ‘to be joyful, cheerful’, clearly a derivative from *šōx* ‘cheerful, joyful’.¹⁹ While *šōxīdan* would yield Modern Persian *šuxidan*, we argue that *šuxidan* ‘to joke’ in our data is more likely to be derived from Modern Persian *šuxi* ‘joke’ rather than from the latter’s predecessor, classical Persian *šōxī* ‘gaiety, impudence’ or from *šōx* ‘joyful’ just mentioned.²⁰

- In our data, *ta’ārofīdan* ‘to kindly offer’ is used with a different meaning than *ta’ārof kardan* ‘to be polite, to compliment, to be deferential, to offer’ (which refers to a tradition of somewhat ritualised politeness). Yet, it seems possible that *ta’ārofīdan* can also be used in the sense of *ta’ārof kardan*.

Two verbs are homophonous with a (clearly entirely unrelated) verb preexisting in Persian:

- *dušidan* ‘to take a shower’ (derived *duš* ‘shower’) is homophonous with *dušidan* ‘to milk’ (also *duxtan*); [p. 28]
- *gāzidan* ‘to accelerate (a car)’²¹ is homophonous with *gāzidan* ‘to bite’.

The followings verbs are found in the encyclopaedic Dehxodā dictionary in the same meaning as in our data:

- *baḥšidan* ‘to have an argument’ (Dehxodā: *nezā’ va jedāl bā zabān kardan* ‘to have a quarrel with the tongue’),

¹⁹ This is not to be confused with Classical Persian *šūxīdan* ‘to soil’, which is derived from *šūx* ‘dirt, impurity’ (cf. also *šūxī* ‘filth, rubbish’).

²⁰ See also Section 2.3 below.

²¹ See also Section 2.3.

- *negāhidan* ‘to look (at); to guard’ (Dehxodā: *negāh kardan, didan; pāsbāni kardan*),
- *gušidan* ‘to listen’ (Dehxodā: *guš kardan; šenidan*),
- *zangidan* ‘to ring (a bell), to call’ (Dehxodā: *āvāz dādan=e zang; zang zadan*),

The speakers whom we interviewed perceive all verbs in Table 4 (including the categories mentioned in this subsection) as ‘new verbs used in informal conversations’, and interpret the corresponding complex predicates as the ‘usual’ form.

We conclude that the verbs in our data, coming from a specific type of source involving younger generations, suggest new formations rather than a continued use.

2.3. Morphological features

While the vast majority of simple verbs in Table 4 are derived directly from the noun by the process discussed in Section 1.1, two types of cases show specific patterns.

Several verbs in Table 4 are derived from an English term, just as one can observe for terms of computer technology and online communication in other languages. This is the case for the items in Table 5.

TABLE 5. Simple verbs in Table 2 and 4 based on English loanwords

Persian				English	
Simple verb		Complex predicate		Verb	Noun
<i>čatidan</i>	‘to chat’	<i>čat</i>	<i>kardan</i>	<i>chat</i>	<i>chat</i>
<i>kilikidan</i>	‘to click’	<i>kilik</i>	<i>kardan</i>	<i>click</i>	<i>click</i>
<i>kopidan</i>	‘to copy’	<i>kopi</i>	<i>kardan</i>	<i>copy</i>	<i>copy</i>
<i>serčidan</i>	‘to search (on web)’	<i>serč</i>	<i>kardan</i>	<i>search</i>	<i>search</i>
<i>imeylidan</i>	‘to send an email’	<i>imeyl</i>	<i>zadan</i>	<i>email</i>	<i>email</i>
<i>gāzidan</i>	‘to accelerate (a car)’	<i>gāz</i>	<i>dādan</i>	--	<i>gas</i>
<i>gugelidan</i>	‘to google’	<i>gugel</i>	<i>kardan</i>	<i>google</i>	
<i>pārkidan</i>	‘to park’	<i>pārk</i>	<i>kardan</i>	<i>park</i>	--
<i>pāsidadan</i>	‘to pass’	<i>pās</i>	<i>kardan</i>	<i>pass</i>	
<i>tāypidan</i>	‘to type’	<i>tāyp</i>	<i>kardan</i>	<i>type</i>	
<i>šutidan</i>	‘to shoot (football)’	<i>šut</i>	<i>kardan</i>	<i>shoot</i>	<i>(shot)</i>
<i>leftidan</i>	‘to leave (an online group)’	<i>left</i>	<i>dādan</i>	<i>(leave)</i>	<i>(leave)</i>

For several items, English presents both a noun and a verb, so that the Persian verbs *čatidan*, *kilikidan*, *kopidan* and *serčidan* could either be wholesale borrowings of the English verbs or derived from the English noun.

However, some items can only be the borrowed verb, viz. *gugelidan*, *tāypidan*, *pārkidan* and *pāsidadan*. Particularly clear is *šutidan*, which can only be based on the verb (*shoot*), not on the noun (*shot*), and *leftidan* (ex. 11), which is even from the past stem of the English verb. This does not prevent it from being used in the present tense in Persian (and the same applies to the complex predicate, e.g., *left mi-d-am* ‘I leave’). In these cases, the corresponding complex predicate employs a non-verbal element that is not used in English, or at least not in this meaning. [p. 29]

Borrowed verbs being integrated into Persian recalls cases such as *fahmidan* ‘to understand’ and *raqšidan* ‘to dance’ in classical Persian, which use Arabic verbal nouns as the verbal base.²² We argue, therefore, that the English verbs are treated as verbal nouns when used in Persian simple verbs and complex predicates in the same way as Arabic verbal nouns have been used for centuries.

²² Lazard (2006: 279) uses the term ‘noms d’action d’origine arabe’.

Some new simple verbs are based on a noun with final vowel, so that the present stem is not identical to the noun it is derived from, namely *šuxidan* ‘to joke’ and *bāzidan* ‘to play’, whose nouns ends in *-i* (*šuxi* ‘joke’, *bāzi* ‘play’), and *qahvidan* ‘to have coffee’ and *‘aṭsidan* ‘to sneeze’, whose noun ends in *-e* (*qahve* ‘coffee’, *‘aṭse* ‘sneezing’).

As shown by the latter two, the morphological description in Section 1.1 is in need of modification: If the nominal (or: verbal noun, see above) ends in *i* or *e*, the vowel is deleted for forming the new simple verbs, thus *qahve* → PRS *qahv-* and *‘aṭse* → *‘aṭs-*, with past stems *qahvid* and *‘aṭsid*. It follows that *bāzidan* and *šuxidan* are likewise derived from *šuxi* and *bāzi*, i.e., these nouns yield the present stems *šux-*, *bāz-* (*mi-šux-am* ‘I joke’, *mi-bāz-am* ‘I play’), which is assigned a past stem in *-id* as any other newly recruited verb.

3. A SOCIOLINGUISTIC STUDY: DATA AND QUESTIONNAIRES

3.1. The data under study

Having discovered the simple verbs in Table 6, we proceeded to a sociolinguistic study.²³ Their acceptability was tested with questionnaires containing sentences featuring one of these verbs, and using the approach described in Section 3.2.

TABLE 6. New simple verbs in Persian used for the sociolinguistic investigation

Simple verb	Meaning	Complex predicate
<i>fekridan</i>	‘to think’	<i>fekr kardan</i>
<i>šoḥbatidan</i>	‘to discuss’	<i>šoḥbat kardan</i>
<i>čatidan</i>	‘to chat’ ²⁴	<i>čat kardan</i>
<i>gušidan</i>	‘to listen’ ²⁵	<i>guš kardan</i>
<i>baḥšidan</i>	‘to have an argument’ ²⁶	<i>baḥš kardan</i>
<i>qahridan</i>	‘to sulk’	<i>qahr kardan</i>
<i>ḥarfidan</i>	‘to talk’	<i>ḥarf zadan</i>
<i>zangidan</i>	‘to call (telephone)’ ²⁶	<i>zang zadan</i>
<i>‘aksidan</i>	‘to take a picture’	<i>‘aks gereftan</i>
<i>darsidan</i>	‘to study’	<i>dars xundan</i>

The reliability of the results was tested with two pseudo-verbs as ‘placebo’, that is, non-existent simple verbs created for the purposes of comparison (Table 7), with examples such as (13).

[p. 30]

TABLE 7. Pseudo-verbs

Pseudo-verb	Intended meaning	(real) Complex predicate
<i>*bāvaridan</i>	‘to believe’	<i>bāvar kardan</i>
<i>*yādidan</i>	‘to learn’	<i>yād gereftan</i>

- (13) **ḥarf=et=o* *bāvar-id-im*
word=PC2SG=RĀ believe-PST-1PL
*intended meaning: ‘We believed what you said (lit. your word).’

²³ We carried out this study in 2019; simple verbs that came to our attention later are included in Table 4.

²⁴ See also Section 2.3.

²⁵ For *gušidan*, see Sections 1.2 and 2.2.

²⁶ See also Section 2.2.

In fact, *qahridan* was initially constructed alongside *bāvaridan* and *yādidan* as a pseudo-verb but it then turned out that it does exist.²⁷ As will be seen in Sections 4.2-4.3, it does indeed yield results like the ‘real’ verbs, which confirms the method chosen.

Conversely, as *qahridan* was initially held to be a pseudo-verb, it was only included in Questionnaire A, not in Questionnaire B (see Section 3.2). Since the results for this verb are thus limited in number, it will only be included in Figures 1 and 2 below.

3.2. The questionnaires and the approach

In preparing the Questionnaires, we constructed sentences based on models such as (1)–(7). Native Persian speakers who did not participate in the test checked all sentences in the questionnaires to make sure they were natural.

The sentences were read to the interviewed speakers in face-to-face meetings or by telephone or WhatsApp. The participants were asked to give grades to the sentences using a Likert rating scale (a method widely employed to measure attitudes or opinions) ranging from 1 point (least acceptable) to 7 points (most acceptable).

Questionnaire A. consists of 45 sentences including one of the verbs in Table 6 or 7. Overall, each element of the categories in Table 8 figures in the Questionnaire for one verb or the other (all persons; negative and positive forms: TAM forms: present indicative, subjunctive, imperative, past tense, imperfect, present and past progressive).

TABLE 8. Categories and terminology used for this paper (example: 1SG forms of *zangidan* ‘to call’)²⁸

Category	Morphology	Structure	Meaning	Examples
Present	<i>mi-zang-am</i>	IPFV-PRS-1SG	‘I call / am calling’	(8a), (11)
Present progressive	<i>dār-am mi-zang-am</i>	have.PRS-1SG IPFV-PRS-1SG	‘I am calling right now’	(3)
Past Imperfect	<i>zangid-am</i>	PST-1SG	‘I called’	(4), (6), (12)
Past progressive	<i>mi-zangid-am</i> <i>dāšt-am mi-zangid-am</i>	IPFV-PST-1SG have.PST-1SG IPFV-PST-1SG	‘I was calling’ ‘I was calling that moment’	(7) (4)
Subjunctive Imperative ²⁹	<i>be-zang-am</i> <i>be-zang-Ø!</i>	SBJV-PRS-1SG IMP-PRS-IMP2SG	‘that I call’ ‘call!’	(1), (2), (5), (11), (14) (11)

[p. 31]

The Questionnaire was answered by 42 respondents between 16 and 58 years of age (Table 9).

TABLE 9. Respondents of Questionnaire A

	16-24 years	25-30 years	31-39 years	40-58 years	Total	
Men	3	5	1	2	11	
Women	14	7	6	4	31	
Sum	17	12	7	6	42	x 45 sentences = 1890 grades

²⁷ We are grateful to David A. S. Moslehi for this information.

²⁸ Examples for the categories perfect, past perfect and infinitive, which were not included in our survey, are given in Section 5.2.

²⁹ In the 2PL, the imperative is identical to the subjunctive.

Questionnaire B was designed to evaluate the influence of the various verbal categories. It includes sentences for each category listed in Table 8 in all persons (positive and negated forms)³⁰ for the verbs in Table 6 (except for *qahridan*, see Section 3.1), that is, 64 sentences for each verb. Owing to the length of this Questionnaire, only 10 respondents agreed to participate (Table 10).

TABLE 10. Respondents of Questionnaire B

	18-24 years	25-30 years	31-35 years	Total	
Men	2	1	2	5	
Women	1	2	2	5	
Sum	3	3	4	10	x 64 sentences x 8 verbs = 5120 grades

After starting the interviews, an additional simple verb was observed, viz. '*aksidan* 'to take a photo'. A set of seven sentences such as (14) was submitted to six respondents (all female, from 21 to 33 years), thus producing 42 grades. As the answers are rather few in number and not diversified enough, the results are only included in Figure 1.

- (14) *az un şafhe=ye ketāb mi-tun-i be-'aks-i?*
 from DEM2 page=EZ book IPFV-be_able.PRS-2SG SBJV-take_photo.PRS-2SG
 'Can you take a picture of that page of the book?' (MES)

3.3. The respondents

The sampling of the respondents was done by a snowball system: starting from personal contacts who agreed to participate, these suggested further participants (relatives or acquaintances) for the study.

The respondents are from various places (Table 11): first of all, Tehran (for all three Questionnaires), then Isfahan and Tabriz (only for Questionnaire A). Two respondents each are from Lar, Shiraz, Lahijan and Ahvaz. Other places are represented by one respondent: Bandar Abbas, Bojnord, Gorgan, Kermanshah, North Khorasan, Lorestan, Zanjan and Sabzevar.

[p. 32]

TABLE 11. Respondents by region

	Questionnaire A	Questionnaire B	Questionnaire ' <i>aksidan</i>	Total
Tehran	9	5	5	19
Isfahan	12	1		13
Tabriz	10			10
Lar	1		1	2
Shiraz	1	1		2
Lahijan	2			2
Ahvaz		2		2
Bandar Abbas	1			1
Bojnord	1			1
Gorgan	1			1
Kermanshah	1			1
Khorasan (North)	1			1
Lorestan	1			1
Zanjan	1			1
Sabzevar		1		1
Sum	42	10	6	58

All interviewed persons belong to middle-class families, are from an urban background and have a higher education; those in student age are studying at university in diverse fields (mainly: chemistry, architecture, computer science or accounting) or are working. All respondents are native speakers of Persian, were born in Iran and have lived in Iran all their life.

4. RESULTS OF THE QUESTIONNAIRES

4.1. Observations and method

The acceptability grades (1 to 7 points, see Section 3.2) which the respondents gave to the sentences of the Questionnaires show a number of statistically significant differences for various parameters. The figures to follow include the results of Questionnaires A and B as described in Section 3.2 above,³¹ totalling a set of 7,010 observations,³² the one in Section 4.2 also includes '*aksidan* (42 additional observations).

For most features to be discussed in what follows, we ran a multivariate regression analysis to make sure that the results do show what we intend to measure, that is, the results are controlled for the variables that one does not intend to measure, and that the composition of the sample does not distort the results.³³ For instance, when measuring the effect of gender, one wishes to exclude that the specific composition of our sample (such as there being more women than men, different composition by age groups, etc.) distorts the results. In his method, testing for statistical significance is part of the analysis.³⁴

[p. 33]

4.2. Results for the individual verbs

The individual verbs have highly different acceptance rates. Figure 1 shows the grades that the respondents gave to the individual verbs.

The verb with the highest acceptance is *čatidan*, obtaining an average of nearly 5 (4.84) out of 7 points, and the least accepted is *gušidan*, with an average of only 2.60. Closely following *čatidan* are *zangidan* and *ħarfıdan*, then '*aksidan* (4.43 points).³⁵ There is a clear difference to the other verbs, of which *darsıdan* and *fekrıdan* show 3.31 points while *baħşıdan* and *şoħbatıdan* (2.79) are less well accepted.

As for the pseudo-verbs (Table 7), **bāvarıdan* and **yādıdan* got significantly lower grades (2.2 points less than any other verb) whereas *qahrıdan* (which was initially held to be a pseudo-verb, see Section 3.1), received a grade similar to that of *şoħbatıdan* and *gušıdan*. This confirms that the method chosen permits to distinguish between real and pseudo-verbs.

³⁰ We did not test negated forms of the progressive (present and past) as this category does not usually occur with negation (see Section 5.2).

³¹ NB that *qahrıdan* is only included in the results presented in Sections 4.2-4.3 (see Section 3.1).

³² 1,890 observations from Questionnaire A; and 5,120 observations from Questionnaire B.

³³ Multivariate regression analysis is a method used in statistics to analyse which variables determine the distribution of values (in our case: the gradings) of the variable (e.g., the specific verb) one is looking at. Technically, we explain the variation in the grades given for a verb by the grammatical attributes of the verb (tense, aspect and mood) while including indicator variables for each respondent (so-called 'fixed effect' model meaning that the grading behaviour specific to a certain respondent is taken into account). In addition, indicator variables for each verb are included as regressors. The coefficients estimated for the verb indicators differ significantly in statistical terms. For more details on regression analysis of linguistic data see Baayen (2008).

³⁴ Statistical significance is expressed by the value of a variable 'p'; when p is below a certain threshold (e.g., 0.05), the probability of the result being by coincidence is so small that one can be reasonably sure that the result is indeed due to the factor one is looking at. – In several of the figures below, a darker shade of blue indicates a higher degree of statistical significance.

³⁵ See Section 3.2 above for '*aksidan*.

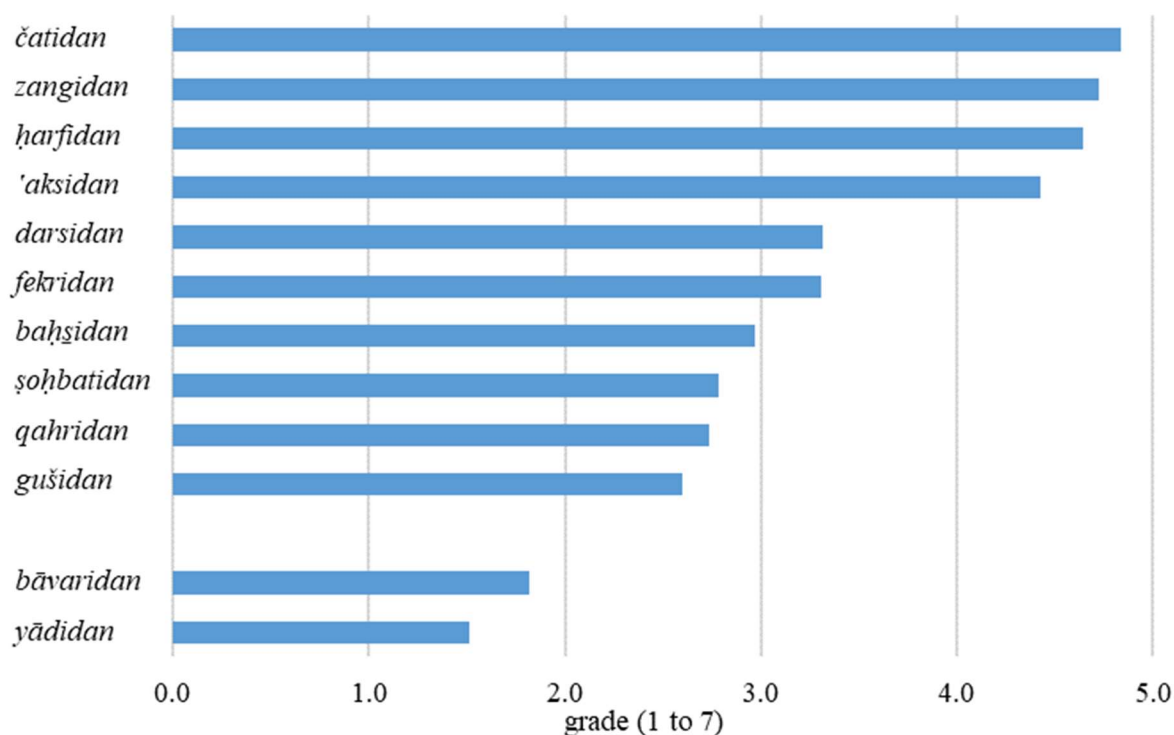


FIGURE 1. Acceptance rate of the individual verbs³⁶

4.3. Results depending on age

The age effect is very interesting. Taking all verbs together, there is a linear function, that is, the younger the respondent, the better the grade, and declining grades as age advances, the correlation being statistically significant for the average of all verbs (the black solid line in Figure 2). The same is true for some individual verbs, viz. *baĥsıdan*, *ħarfıdan* and *qahrıdan*, where acceptance declines continually with age.

The other verbs do not pattern in this way, though. Three verbs show a curve with a peak: *ŝoĥbatıdan* shows a curve with a peak at 39 years, decreasing on both sides. It seems to be [p. 34] specific to a certain generation, and acceptance declines for respondents who are older, and also for young people who are maybe more influenced by what they learn at school or university. However, the two the best accepted verbs (see Figure 1), viz. *čatıdan* and *zangıdan*, show only a very slight rise and a decline after the peak, which is at around 30 years.

Quite conversely, the other verbs have an inverted curve: *fekrıdan*, *guŝıdan* and *darsıdan* start out with more or less the same good acceptance among young people as *čatıdan* and *zangıdan*, but this is followed by a marked decline; *guŝıdan* even has a pronounced minimum, confirming some speakers' impression (see Section 1.2) that *fekrıdan*, *guŝıdan* and *darsıdan* are more commonly used by teenagers. All five verbs then have roughly the same grades with the eldest respondents in the sample.

³⁶ The figure shows the simple mean of the grade. The result is almost exactly the same if one runs a regression analysis controlling for gender, age, location and grammatical form.

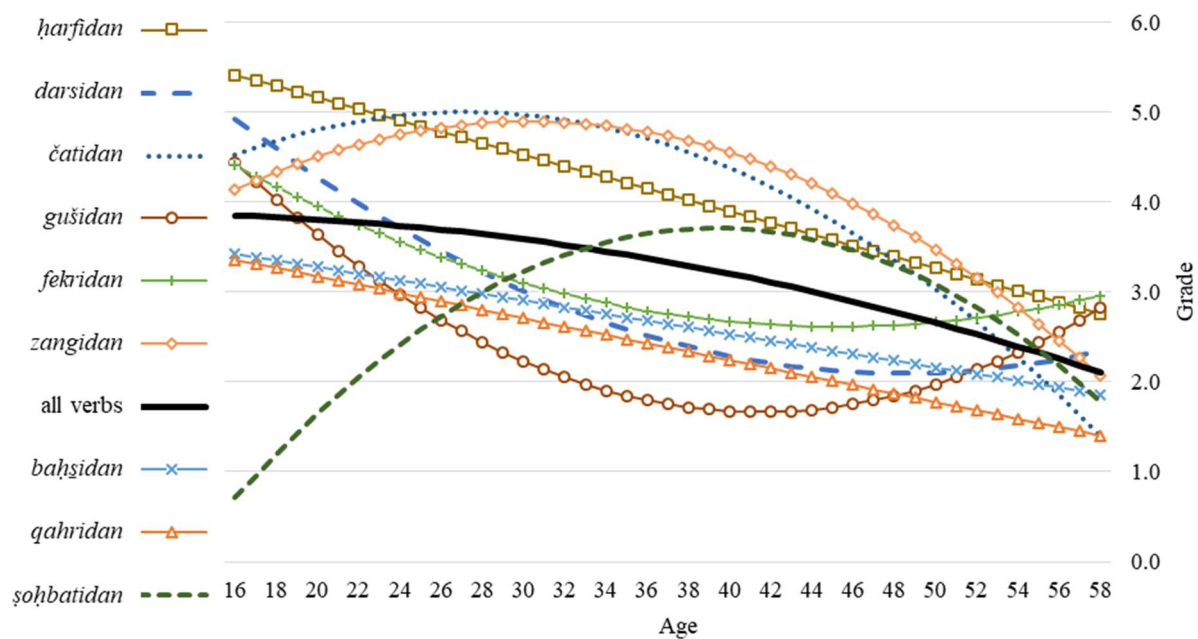


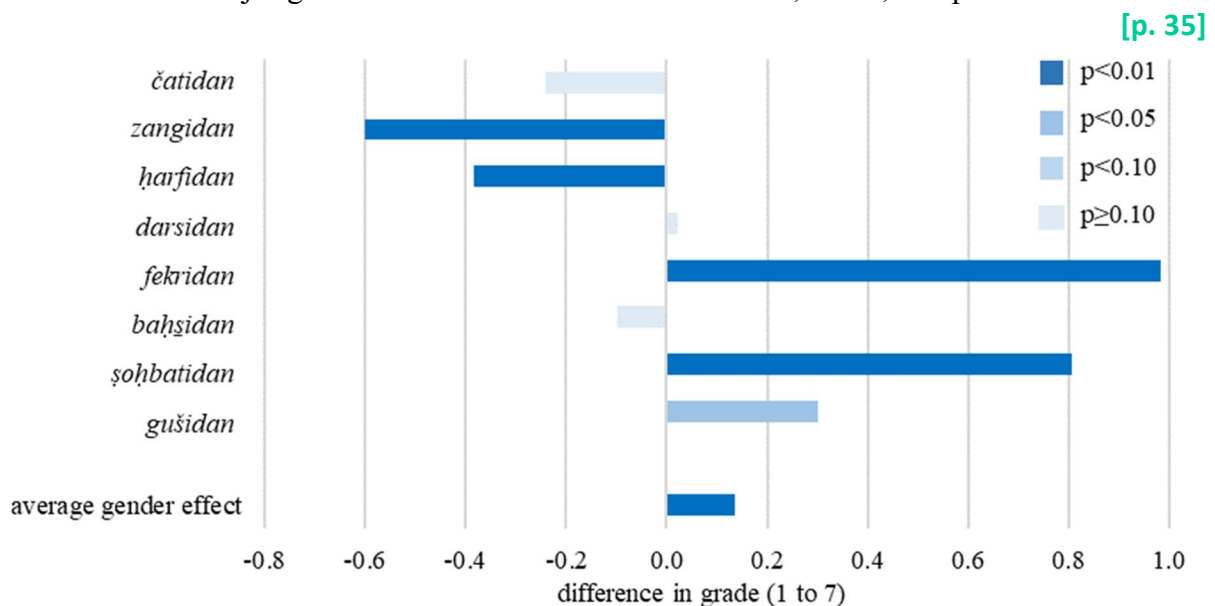
FIGURE 2. Acceptance rate of the individual verbs depending on age³⁷

4.4. Results depending on gender

The gender difference is not very marked, but statistically significant. When one ‘calculates out’ the somewhat differing age of male and female respondents, men give slightly higher marks than women by 0.18 points; if one omits the pseudo verbs by 0.14.

Again, there is an important variation for the individual verbs. Figure 3 shows the gender difference in acceptance for the individual verbs. Compared with women, men show a lower acceptance of *zangidan* and *harfidan*, whereas the ratings for *şohbatidan* and *fekridan* are substantially higher.

There are no major gender differences for the various tense, mood, and person forms.



[p. 35]

³⁷ Schematic presentation. The curves are derived from regressing the grades given to a verb on the age of the respondent and the squared term of the age. In case both estimated coefficients of age and age² were statistically significant, the inflexion point and the slope of the curve were estimated. The list on the left corresponds to the order of the curves on the left side of the figure.

FIGURE 3. Gender differences in the acceptance rate of the individual verbs: grades given by men compared to grades by women.

The table shows differences in grades given by men compared to grades given by women (controlling for age), who, being more numerous among our respondents, are set as the point of comparison at ‘0.0’. In the case of *fekridan*, for example, men give on average a grade higher by almost 1.0 points than women. The darker the blue, the higher the statistical significance (p having a low value).

4.5. Results depending on region

There is no significant difference between the grades of respondents from Tehran and those from all other regions taken together.

However, looking at the average is misleading, as in fact the various regions pattern very differently. As shown by Figure 4, Tabriz yields nearly the same results as Tehran while speakers from Ahvaz and Sabzevar give significantly higher grades. Grades from Isfahan, Shiraz and the remaining regions (see the list in Table 11) are much lower.

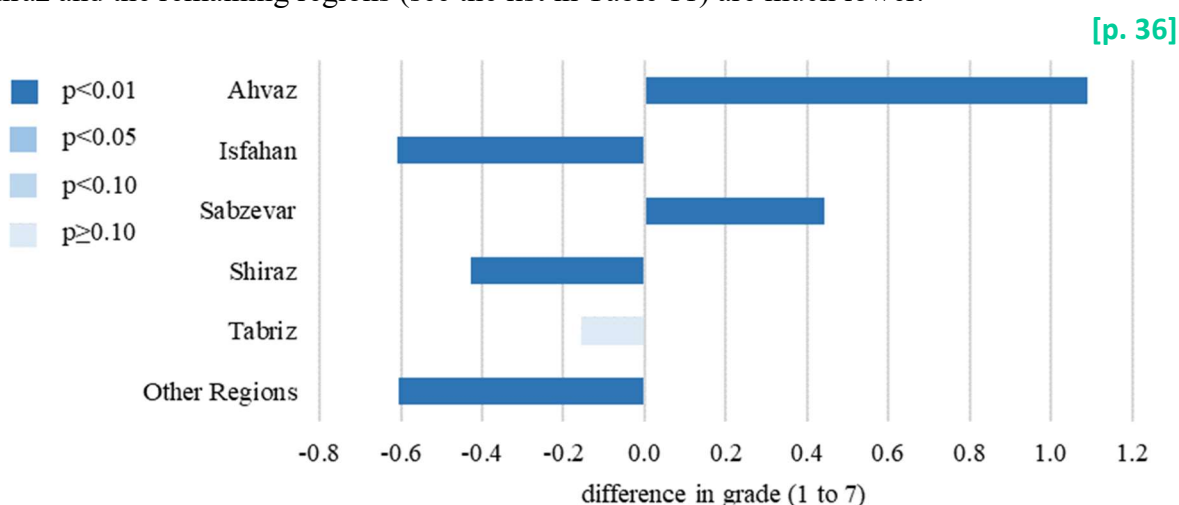


FIGURE 4. Regional differences in acceptance rates: grades relative to Tehran.³⁸

‘Other regions’ include respondents from Bandar Abbas, Bojnord, Gorgan, Kermanshah, North Khorasan, Lahijan, Lar, Lorestan and Zanjan.

Figure 5 shows the acceptance of the individual verbs by respondents from different regions. The lighter the colour, the lower the difference in grading to speakers from Tehran. Noteworthy results include that in spite of rating all verbs taken together lower, Isfahani respondents give rather good rates for *fekridan*. Conversely, *gušidan* gets lower acceptance in Tehran than in all other places. Respondents from Shiraz like *baħsidan* although they are on the whole not very fond of the new verbs.

	<i>čatidan</i>	<i>zangidan</i>	<i>ħarfidan</i>	<i>darsidan</i>	<i>fekridan</i>	<i>baħsidan</i>	<i>šoħbatidan</i>	<i>gušidan</i>
Ahvaz	0.70	-0.45	0.81	1.85	1.86	1.36	0.03	2.39
Isfahan	-1.43	-0.86	-1.53	-2.20	0.82	-1.24	0.07	1.89
Sabzevar	1.14	1.09	0.42	-1.05	-0.42	0.91	0.59	0.73
Shiraz	-2.33	-0.07	-1.43	0.37	0.20	1.08	-1.37	0.05
Tabriz	-1.17	0.10	0.32	0.09	0.35	-0.56	-0.60	1.16
Other Regions	-0.76	-0.56	-0.63	-0.69	-0.48	-1.05	-0.84	1.07

FIGURE 5. Regional differences in acceptance rates of verbs: grades relative to Tehran.³⁸

The table shows differences in grades compared to respondents from Tehran. – Lower relative grades (negative values) are shown in red, higher relative grades (positive values) in blue.

³⁸ Results of multivariate regression analysis (controlling for age and gender).

4.6. Results depending on verb form

More to the linguistic side of the topic, acceptance rates differ significantly depending on tense, aspect, mood, and other categories of the individual verb form. Figure 6 shows the results from the two Questionnaires for the TAM categories. The imperfect (type *mi-zangid-am* IPFV-call.PST-1SG ‘I was calling’, see Table 8) is the least accepted category and serves as point of comparison. The past and the imperative are better rated without this being statistically significant. The present gets very high grades, but the subjunctive and the progressive (both present and past) are rated significantly better than the imperfect as well, with the present progressive even better rated than the present tense.

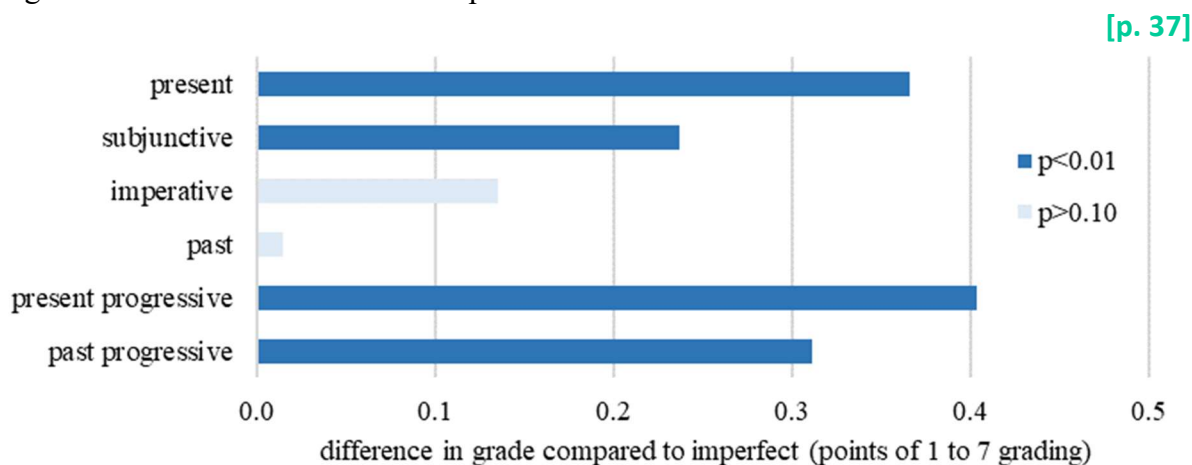


FIGURE 6. Acceptance rates of simple verbs by tense and mood.³⁹

Interpretation: A verb form in present tense receives on average a 0.39 better grading than a verb in the imperfect.

While Figure 6 is from the viewpoint of the imperfect (as the least accepted category), showing which categories are rated insignificantly or significantly better, Figure 7 shows the results for the various categories compared against each other. For instance, from the viewpoint of the present (first item in the list on the left), the difference to the progressives is not statistically significant (indicated by ‘0’ in the same line) while its ratings are much better than those of the past and imperfect (+++), moderately better than the imperative (++) and slightly better (+) than the subjunctive.

	present	subjunctive	imperative	past	imperfect	present progr.	past progr.
present		+	++	+++	+++	0	0
subjunctive	-		0	+++	+++	-	0
imperative	--	0		0	0	--	-
past	---	---	0		0	---	---
imperfect	---	---	0	0		---	---
present progr.	0	+	++	+++	+++		0
past progr.	0	0	+	+++	+++	0	

FIGURE 7. Differences between tenses and moods in the acceptance of simple verbs.³⁹

+++ , ++ , + ; --- , -- , - : a form of the category present, subjunctive etc. listed vertically on the left is rated significantly better (plus) or lower (minus) than a form of another category among those listed horizontally in the first row with a statistical error of less than 1%, 5%, and 10% respectively. ‘0’ indicates that there is no statistically significant difference in the rating of the respective pair of forms. Everything indicated by one or more + or - in this figure is a statistically significant difference.

Evaluating the acceptance rates of verbs for different TAM forms, Figure 8 demonstrates that ratings differ most strongly for the imperative. For instance, if the verbs *ħarfıdan* and

³⁹ Results of multivariate regression analysis (controlling for age, gender and verb).

fekridan are presented in the imperative, the grading is particularly low, while *guşidan* receives the highest grading when it is presented in the imperative. For present and past, acceptance rates vary more across individual verbs than for present progressive and past progressive.

[p. 38]

	<i>çatidan</i>	<i>zangidan</i>	<i>harfidan</i>	<i>darsidan</i>	<i>fekridan</i>	<i>bahşidan</i>	<i>şoşbatidan</i>	<i>guşidan</i>
PRS	0.37	0.47	0.46	-0.44	-0.55	-0.35	-0.24	0.05
SBJV	-0.26	0.96	-0.07	-0.11	0.01	-0.11	0.25	0.16
IMP	-0.70	0.41	-1.08	-0.71	-1.13	-0.73	-0.76	0.47
PST	-0.33	0.69	-0.55	-0.61	-0.73	0.05	-0.61	-0.10
PROG.PRS	-0.14	0.56	-0.31	0.08	-0.47	-0.47	-0.16	0.18
PROG.PST	-0.61	0.43	-0.10	-0.27	-0.57	0.20	-0.48	0.22

FIGURE 8. Differences in acceptance rates of each verb depending on tense and mood.⁴⁰

The table shows differences in grades for a specific TAM form of each verb (relative to the grades for the imperfect) compared to the values shown in Figure 6, that is, compared with the average grade for each category of all verbs together (relative to the imperfect). – Lower relative grades (negative values) are shown in red, higher relative grades (positive values) in blue.

As far as other verbal categories are concerned, negated verb forms get far lower grades than the positive ones for all categories.

For number, the singular is better accepted than the plural, but the difference is not statistically significant.

The third person is the least accepted of the persons; the better rating of the 1st and 2nd person each is statistically significant (Figure 9).

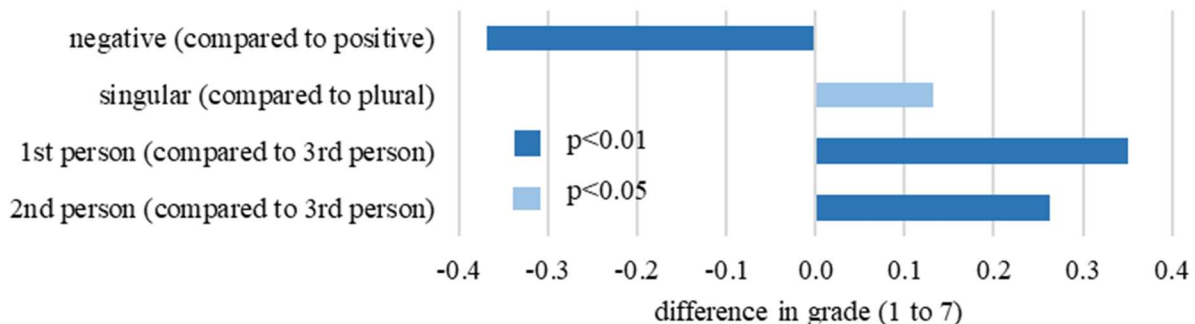


FIGURE 9. Acceptance of simple verbs by negative/positive, number and person.⁴¹

Interpretation: A negated verb form receives on average a 0.38 lower grading than a form in the positive.

The ranking of persons in the different TAM categories yields interesting results (Figure 10). Overall, the 2SG is the best accepted person, followed by the 2PL, 1PL, and 1SG, with the 3PL being last. The good ranking of the 2SG applies throughout the TAM categories, while the 1SG and the 1/2PL show a low ranking in one or the other category. In most categories, the difference in ranking of the person (i.e., the difference from the average for this category as shown in Figure 6) is statistically significant, with the exception of the imperfect and the progressive present, where the difference between the highest and lowest grading is small (the grades for the imperfect are nearly the same for all persons).

[p. 39]

Looking again at the behaviour of the individual verbs, differences in acceptance rates of negative/positive, number and person are less marked than for tense and mood (Figure 11). One exception is the singular, which receives significantly higher grades for *zangidan* and significantly lower ones for *bahşidan*. 1st and 2nd person show a much higher acceptance for *harfidan* as for any other verb, while the negative of the same verb receives very low grades.

⁴⁰ Results of multivariate regression analysis (controlling for age, gender, negative/positive, number and person).

⁴¹ Results of multivariate regression analysis (controlling for age, gender and verb).

	PRS	SBJV	PST	IMPF	PROG.PRS	PROG.PST	total
1SG	6	1	4	4	5	3	4
2SG	1	2	1	2	1	4	1
3SG	3	6	5	5	4	1	5
1PL	5	3	2	3	3	5	3
2PL	2	5	3	6	2	2	2
3PL	4	4	6	1	6	6	6
significance	*	*	*			*	*

FIGURE 10. Ranking depending on person and TAM form.⁴²

1 (darkest background) = highest rank; 6 (white background) = lowest rank. * = The differences between the persons in this TAM category are statistically significant.

Interpretation: Among the forms of the subjunctive of all verbs taken together, the 1SG receives the highest grades, followed by the 2SG and the 1PL; the 3SG is ranked lowest and the 3PL and 2PL in between.

	<i>čatidan</i>	<i>zangidan</i>	<i>ħarfidan</i>	<i>darsidan</i>	<i>fekridan</i>	<i>baħšidan</i>	<i>šoħbatidan</i>	<i>gušidan</i>
NEG (compared to positive)	0.02	0.15	-0.40	0.19	-0.12	0.28	0.27	0.12
SG (compared to PL)	0.15	0.63	-0.02	-0.09	-0.04	-0.57	-0.14	-0.06
1st (compared to 3rd)	0.09	0.24	0.60	0.31	-0.39	-0.12	-0.37	-0.35
2nd (compared to 3rd)	0.13	0.18	0.57	0.13	0.02	-0.04	-0.06	-0.32

FIGURE 11. Differences in acceptance rates of negative/positive, number and person for each verb.⁴³

The table shows differences in grades for a specific verb form depending on the verb as compared with the values shown in Figure 9, that is, the average grade for the respective category across all verbs. – Lower relative grades (negative values) are shown in red, higher relative grades (positive values) in blue.

5. DISCUSSION AND CONCLUSION

It has been widely argued that simple verbs are not productive in modern Persian, and that new verbal expressions are only formed by combining a non-verbal element with a light verb. By analysing data from social media and evaluating its acceptability through questionnaires, our objective was to demonstrate that this claim is no longer valid. In present-day colloquial Persian, a number of new simple verbs have been created and are fairly widely accepted. This basic point is both unexpected and important. Additionally, our study yielded the following findings.

[p. 40]

5.1. Results for various sectors of society

Concerning age, taking together all verbs in our study, acceptance of the simple verbs declines with age, that is, the younger the respondent, the better the grade. This agrees with speakers' judgment (see Section 1.2) that some of them are chiefly used by teenagers.

However, when one looks at the individual verbs, there is considerable variation, implying that age does not always have a negative impact for the acceptability. Most verbs have their specific pattern, for example, a peak or a low at a certain point of the curve. Some verbs seem to be specific for a certain generation while both younger and older respondents do not readily accept them. Particularly interesting are the verbs which are rated rather high by both young and older respondents while the middle generation in our sample (30–45 years) does not appreciate them. It is not clear what motivates the specific distribution of the individual verbs.

The significantly lower results of the pseudo-verbs demonstrate the relevance of such a 'dummy' test and confirm that the other verbs are taken for real by the respondents.

⁴² The imperative is not included in this figure as it only exists in the 2nd person (see Table 8).

⁴³ Results of multivariate regression analysis (controlling for age, gender and TAM form).

Concerning gender, men turn out to accept the simple verbs somewhat more easily than women. Again, the variation among the individual verbs is important, some of them being significantly better accepted by one or the other gender.

This is interesting in the light of Labov's findings (Labov 1990) concluding that, while men's speech shows a higher frequency of nonstandard forms, women's speech shows a higher frequency of incoming (newly innovated forms). It would be interesting to study whether the somewhat better acceptance by our male respondents could be seen in the context of men preferring working class registers (see, e.g., Trudgill 1972). If this explanation is on the right track, the men in our sample would perceive the new simple verbs as belonging to the speech of a sector of society associated with lower prestige,⁴⁴ and this perception would be stronger than the women's perception of these verbs as incoming forms.

It is noteworthy, though, that the acceptance rates in our study do not differ very much depending on gender (although the difference is statistically significant) and again, it is not clear why they differ so much depending on the individual verb.

Geographical origin patterns somewhat similarly to gender in the sense that all verbs taken together, the difference between Tehran on the one hand and all other regions on the other is not very marked – in this case, it is not even statistically significant – but the origin does yield widely differing results for the individual verbs, respondents from this or that region rating one or the other verb significantly higher or lower than respondents from Tehran.

Importantly, though, all verbs taken together, it is not so that respondents from Tehran would rate the innovated verbs better than those from other regions. This shows that the phenomenon of the new simple verbs is found all over Iran and is not limited to the colloquial speech of the capital of Iran, where one might expect linguistic change to originate.

Conversely, the formation of new simple verbs does not seem to be found in Tajiki or Dari (Afghan Persian). Quite to the contrary, the latter has a tendency of replacing even those inherited simple verbs that continue to be used in Iranian Persian by complex predicates employing a non-finite form of the verb (*porsidan* → *porsān* (PTCP.PRS) *kardan* 'to ask', *bastan* → *bastā* (PTCP.PRF) *kardan* 'to bind', etc.).⁴⁵ We thus seem to be dealing with a phenomenon belonging to the grammar of Iranian Persian.

[p. 41]

5.2. The evaluation of various verb forms

Concerning the individual verb forms, the singular is rated slightly better than the plural, but the difference is not statistically significant. The 1st and 2nd person are rated better than the 3rd. This contrasts with the 3rd singular often being the least marked form cross-linguistically – so much so that new paradigms may be based on it, as is the case for Persian *hast-am* and Polish *jest-em* etc., both historically 'exist.3SG-1SG', but synchronically 'I am', that is, a verb form inflected for the 3SG is used as stem for a new paradigm (Watkins 1962: 90–96). One might thus expect that it is best accepted as the base of the innovated paradigm, all the more as it is zero-marked in the past, thus morphologically particularly simple (see Table 3). The fact that the 3rd persons are rated lowest is also noteworthy since the 3rd plural is used for impersonal statements (e.g., 'they say' corresponding to English *it is said*) and the 3rd singular for an undetermined person (Lazard 2006: 125–126), which should add to their frequency, thus potentially influencing the acceptability. Perhaps there are pragmatic factors that motivate the better acceptance of the 1st and 2nd person.

Negated forms are rated lower than positive ones. Maybe this can be seen in the negative forms being cognitively more complex than affirmative ones (Miestamo 2005: 196), thus

⁴⁴ Note that all participants in our study are rather well-educated (see Section 3.3).

⁴⁵ We are indebted to Lutz Rzehak (Berlin) for this information.

possibly less easily accepted in new verbs. Also, and although not the same phenomenon, the distribution of negated forms is often more restricted than that of the affirmative ones, and in many languages ‘fewer grammatical distinctions are available in the negative than in the affirmative’ (Miestamo & van der Auwera 2011: 68). Negated forms may be restricted in terms of aspect, morphological categories, semantics, or discourse contexts (Miestamo 2005, see also Næss 2007: 41–42, 114–122). A case in point is the Persian progressive (*dār-am nāme mi-nevis-am* ‘I am writing a letter (just now)’), see Table 8), which is not normally negated:⁴⁶ the action of ‘I am not writing in this moment’ may be seen as amounting to ‘I am not writing / I don’t write’, so that marking such a form as progressive would be pointless.⁴⁷

Concerning the TAM forms, the present progressive just mentioned gets the highest values, closely followed by the simple present and the past progressive. This is noteworthy since the progressive contains two inflected verb forms (see Table 8), which could be seen as cognitively more difficult to process than a category with one inflectional word form. Also, the past progressive (*dāšt-am mi-nevešt-am* have.PST-1SG IPFV-write.PST-1SG ‘I was writing in that moment’) contains the imperfect (*mi-nevešt-am* ‘I was writing’), which is the least accepted category among all TAM forms.

The category which is by far closest to the imperfect in terms of low gradings is the simple past, in spite of this category being very frequent and also morphologically the simplest among those studied here.

Another interesting difference is the one between the subjunctive (rather well accepted) and the imperative (acceptance rather low) in spite of both being formed with the aspectual prefix *be-* (with variants), and the 2PL being identical for both categories.

It is not clear what motivates the differences in acceptance of verb forms, some of which seem counter-intuitive. Clearly these results encourage further research on the acceptance of the various verbal categories in innovated verb forms cross-linguistically.

As with all other parameters, there are important differences between the individual verbs regarding the acceptance of various TAM forms. [p. 42]

All the forms we tested are inflected. It would be interesting to compare them to categories that use a participle or the infinitive.⁴⁸ We came across examples such as (10), (15), (16) containing the perfect participle, and the infinitive is also found in our data (17).

(15) *ḥālat=e nešastan=et ‘avazid-e*
 perfect manner=EZ sit.INF=PC2SG change.PST-PRF(3SG)
 ‘Your way of sitting has changed.’ (WEB)

(16) *man faqaṭ qāleb=am=o faqaṭ ‘avazid-e bud-am*
 past perfect I only template=PC1SG=RĀ only change.PST-PRF be.PST-1SG
 ‘I had only changed my template.’ (WEB)

(17) *ba’d_ az tamumidan=e māh=e ‘Asal,*
 infinitive after complete.INF=EZ month=EZ PN
ye se čahār ruz=i afsordegi=ye mozmen dāšt-am
 one three four day=IDV depression=EZ severe have.PST-1SG
 ‘After the month Asal ended I had a severe depression for three-four days.’ (WEB)

⁴⁶ In fact, negated forms do occasionally occur, although they are widely regarded as ungrammatical. See Vafaeian (2018) for a study of the Persian progressive (pp. 79–118) and its restriction on negation (pp. 136–138), and for progressives cross-linguistically in general. The statement that ‘[n]o asymmetry is found’ in Persian (Miestamo 2005: 336, 409), that is, negated patterns correspond to affirmative ones, is thus not correct when one includes the progressive.

⁴⁷ We did not test the negated progressive in our study.

⁴⁸ Many modal constructions in Persian use the subjunctive with an auxiliary or an impersonal construction and do not fall within this category.

5.3. The verbs in our study

As mentioned in 5.1-5.2, the individual verbs pattern quite differently for every parameter we tested, be it the features of the respondents (age, gender, regional origin) or those of the verb form. It is not clear what factors this depends on.

It does not seem to depend on the origin of the verb, though. A systematic study remains to be carried out, but *čatidan* ‘to chat’, which is the only verb in our sociolinguistic study that is borrowed, rather seems to suggest that the origin does not play a role for the acceptance of the verb. *čatidan* is the best accepted verb in our study (see Figure 1), but it is closely followed by *zangidan* ‘to call’ and *ħarfidan* ‘to talk’, which are based on Persian *zang* and Arabic *ħarf*, respectively, both of which have since long existed in the Persian lexicon.

When one ‘zooms into’ Figure 1 to compare *čatidan* against *zangidan* plus *ħarfidan* together (Figure 12),⁴⁹ it turns out that *čatidan* is rated better than the other two verbs by each group (women, men, ages below and above 30, participants from Tehran and from other regions), but it is only the group of male participants where the difference is statistically significant, that is, where the average grades for *čatidan* are significantly higher than those of *zangidan* and *ħarfidan*.

For all other groups, the differences are not statistically significant.

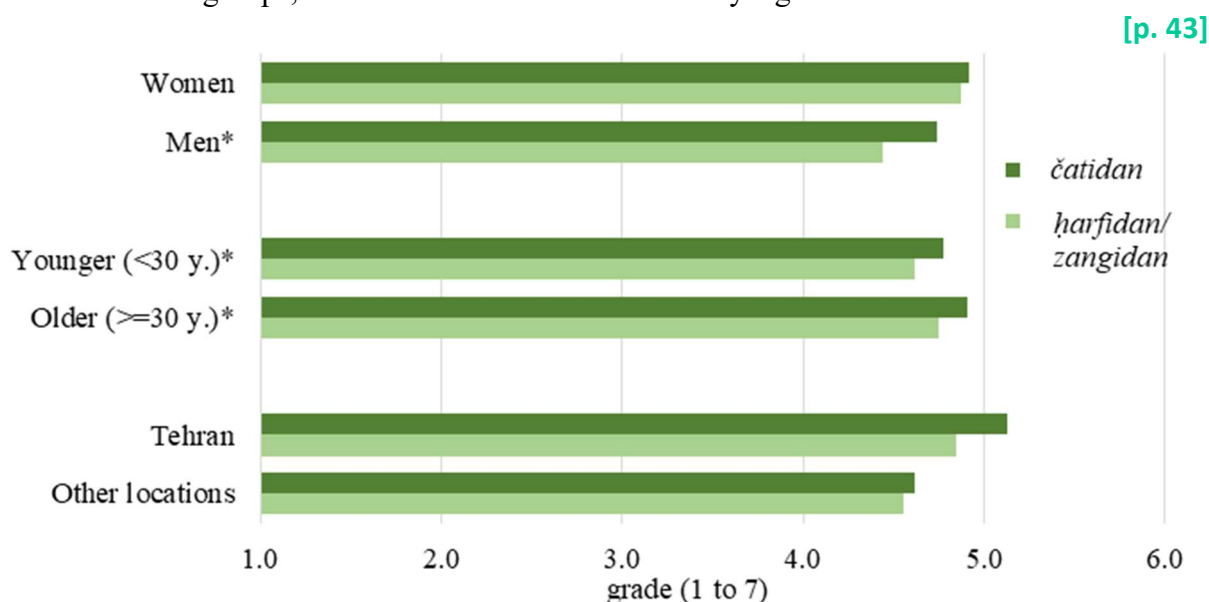


FIGURE 12. Average rating of *čatidan* vs. *ħarfidan* and *zangidan* by gender, age and region.

* = The difference between the two groups of verbs is statistically significant.

Interpretation: The average rating for *čatidan* by women is 4.91, and for *ħarfidan* and *zangidan* 4.87. The difference is not statistically significant.

In terms of detailed age-related acceptance, *čatidan* is particularly close to *zangidan* (see Figure 2), both verbs showing a curve with a maximum at ca. 30 years, while *ħarfidan* has a straight line. On the whole, then, *čatidan* does not pattern significantly differently from similarly well accepted verbs which are based on lexical items already existing in the language.

As shown in Table 5, simple verbs based on English borrowings frequently occur in the field of computer and internet technology. However, one cannot say that innovated simple verbs are specific to borrowings: most items in Table 4 are not based on recently borrowed lexical items, that is, the new simple verbs are etymologically diverse.

⁴⁹ It would not make much sense for the question of borrowings to compare *čatidan* to, say, *fekridan* or *darsidan* as these have rather different rates anyway, so that one would not know whether it is indeed the borrowed-ness that causes the difference.

Similarly, verbs referring to communication, particularly online communication, have a significant share among the simple verbs that came to our attention, but Table 4 shows again that there are also a large number of verbs from other semantic fields.

Indeed, since our results suggest the productivity of simple verb formation (see Section 5.4), it seems only natural that new activities arising by new technological means are represented among the new simple verbs. Of course English-derived verbs related to the use of computers are found in many languages, cf. German verbs such as *klicken*, *chatten*, *emailen*, *googeln*, *uploaden*.⁵⁰ Still, the interest here is that the derivation of new simple verbs has been said to be non-productive or inexistent in Persian while the derivation of verbs directly from nouns is a productive process in German (just as in a number of other languages).

The semantic fields of new simple verbs would be an interesting topic for future research, and it would also be interesting to investigate the type of nominal the simple verb is based on, such as statives vs. eventives, or the telicity of the simple verbs. [p. 44]

Another question one might ask is whether there is a difference in acceptance depending on the light verb that the corresponding complex predicate is formed with. Considering that among the new simple verbs we found those whose corresponding complex predicate shows *kardan* are by far the most frequent (see Section 2.1), we compared those in our sociolinguistic study (viz. *čatidan*, *fekridan*, *baḥṣidan*, *šoḥbatidan*, *qahridan*, *gušidan*, see Table 6) with the remaining ones taken together (viz. *zangidan*, *ḥarfidan*, *‘aksidan*, *darsidan*, that is, the simple verb equivalents of *zang zadan*, *ḥarf zadan*, *‘aks gereftan*, *dars xundan*).

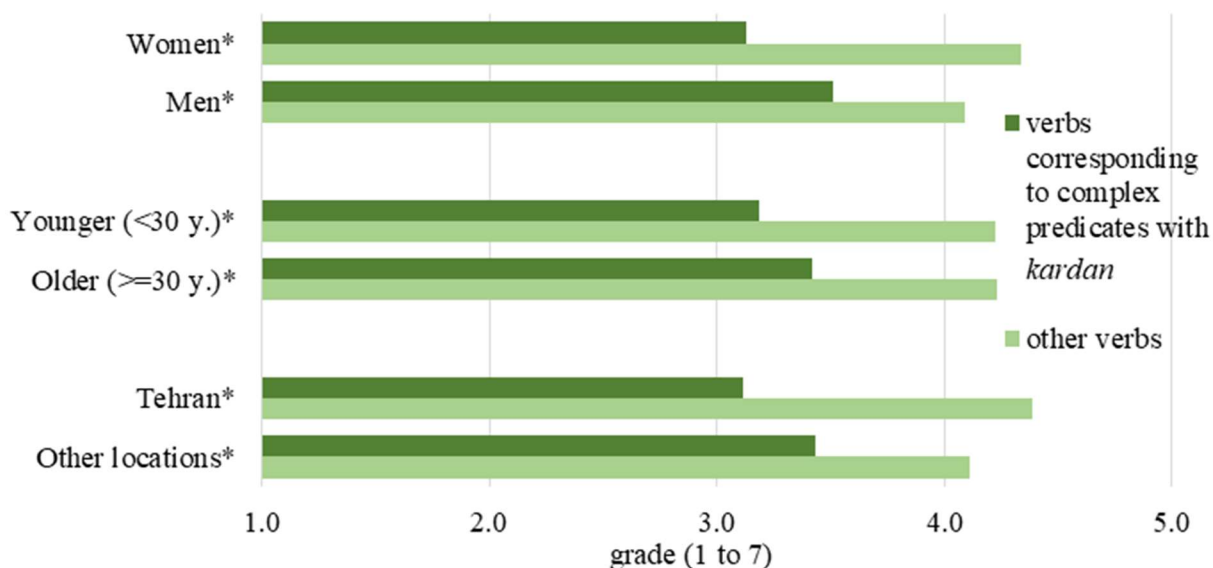


FIGURE 13. Average rating of simple verbs by light verb of corresponding complex predicate.

* = The difference between the two groups of verbs is statistically significant.

Interpretation: The average rating of verbs that correspond to complex predicates with *kardan* by women is 3.1, whereas women give an average rating of 4.3 to all other verbs. The difference is statistically significant.

Interestingly, the *kardan*-corresponding items are less well accepted, with a significant difference for all groups of respondents (Figure 13). This is particularly so for our female participants when compared to men and for the speakers from Tehran as opposed to those from other regions.

⁵⁰ Note that German also borrows the English orthography (*emailen*, *uploaden*) and pronunciation (the diphthongs *ei* and *ou* do not exist in German), but it does adjust verbs to German morphology. For instance, *googeln* ‘to google’ (cf. the entry in the authoritative dictionary of Standard German DUDEN <https://www.duden.de/rechtschreibung/googeln>) shows *-eln* (instead of **googlen*), which is due to an existing class of verbs in *-eln* (among them ‘diminutives’ such as *brutzeln* ‘to fry (*braten*) in a non-serious way’) and *-ern* (*wandern* ‘to hike’, etc.). Borrowed verbs containing a preverb are classed with the inherited separable verbs (*uploaden* → PRS.1SG *ich lade up*, PRF *up-ge-load-et*).

Indeed, when looking at Figure 1, it springs to mind that *čatidan*, the best accepted verb (corresponding to *čat kardan*) is followed by all the verbs whose complex predicates do not show *kardan*, and the remaining *kardan*-items rank lower on this scale. These results are noteworthy given the relative frequency of simple verbs corresponding to a complex predicate with *kardan* ‘to do’.

5.4. Conclusion

5.4.1. Productivity and grammaticalisation

One question that arises for future research is whether a new simple verb being formed depends in some way on an existing complex predicate using the same nominal, or if the derivation of a simple verb has nothing to do with the existence of such a complex predicate.

In this context, the high percentage of simple verbs in Table 4 corresponding to complex predicates with *kardan* springs to mind. Historically as well, *kardan* is certainly the best established light verb already in Middle Persian, and (as argued in Korn 2013) probably was [p. 45] the earliest one to arise. In Middle Persian, denominatives such as *nāmīdan* ‘to call (by a name)’ tend to correspond to noun-verb combinations (*nām kardan* ‘id.’). They often show a formation with *-ēn-* besides (e.g., *nām-ēn-īdan* ‘id.’).⁵¹

Conversely, our list of new simple verbs also includes instances whose complex predicate shows another light verb. And complex predicates with light verbs other than *kardan* are not a recent or accidental phenomenon. This is demonstrated by (classical) Armenian, which shows complex predicates calqued on Iranian models. These show, besides ‘do’, the Armenian equivalents of *zadan* ‘hit’, *dādan* ‘give’ and *dāštan* ‘have/hold’, demonstrating that these have been in use in the spoken (or colloquial) language as light verbs since Early Middle Iranian (Korn 2013: 32, 48–53).

Moreover, as seen in Section 5.3, new simple verbs corresponding to complex predicates with *kardan* are not better accepted than those corresponding to complex predicates with other light verbs, that is, the existence of a complex predicate with *kardan* does not seem to increase the acceptability of a new simple verb, rather to the contrary. Maybe this rather speaks in favour of the formation of simple verbs being independent of the complex predicates, but one should perhaps not draw too far-reaching conclusions from our sample. It still is so that new simple verbs are on the whole well accepted.

A study of simple verbs throughout the history of Persian remains to be undertaken, but, as mentioned in Section 1.1, the productive way of forming new verbs is the creation of complex predicates, and the derivation of simple verbs is a marginal phenomenon – so much so as to motivate numerous researchers to state its non-existence in Modern Persian.

At the same time, it seems that the possibility of forming new simple verbs never entirely ceased to exist. To what extent it was productive at which period, and how frequent and accepted such verbs were, is not known. It is possible that the formation of simple verbs has been a process limited to (or more typical of) the spoken language, which would explain the impression – necessarily based on written sources for the premodern periods – that it was or is (more or less) inexistent.

It has been pointed out (see Section 1.1) that once a language has acquired complex predicates, this pattern readily integrates borrowed material. This also applies to the new simple verbs (see Section 2.3), for example, *šutidan* ‘to shoot’ (besides *šut kardan*). Particularly interesting is *leftidan* ‘to leave (e.g., an internet forum)’ (besides *left dādan*), which is from the past stem of the English verb, but used as a present stem in Persian (*mi-left-am* besides *left mi-*

⁵¹ The suffix *-ēn-* also derives causatives (e.g., *dān-* ‘to know’ → *dān-ēn-* ‘to cause to know’); the function of this suffix is thus essentially transitive, agreeing with that of complex predicates with *kardan* (cf. Korn 2013: 31f, 45f).

d-am ‘I leave’). The derivation of Persian verbs from borrowed verbs is reminiscent of cases such as *fahmidan* ‘to understand’ and *raqsidan* ‘to dance’, which are derived from Arabic verbal nouns (*fahm* ‘understanding’, *raqs* ‘dancing’). This suggests that the English verbs are treated as verbal nouns when they are integrated into Persian.

5.4.2. Synthetic and analytic structures

While Butt 2010 stresses the difference between grammaticalisation proper (as seen, e.g., in the development of auxiliaries) and the rise of complex predicates, we argue that there is a parallelism in the sense that in both cases, analytic structures replace inherited synthetic ones, for example, synthetic passives or perfects being replaced by analytic patterns and simple verbs giving way to complex predicates.⁵² [p. 46]

One would expect that this development is unidirectional in the same way as, or parallel to, grammaticalisation processes have been assumed to be. In this perspective, the emergence of new simple verbs is unexpected, implying as it does a return to a synthetic structure. Indeed, many linguists assume a cycle in which languages change from ‘isolating to agglutinating, to fusional, back to isolating, and so on’ (Dixon 1997: 42); and ‘[s]ynthetic languages such as Old English change into more analytic languages’ (van Gelderen 2011: 345).⁵³ Grammaticalisation can be seen as a phenomenon occurring in this cycle, which is seen as unidirectional (thus, e.g., van Gelderen (2011: 7)). Therefore, one might not expect that the tendency to analytical structures should be reversed.

It could be so that to the extent that there is a certain increase productivity of the derivation of verbs from nouns, this is limited to the speech of a younger generation and that its use is particularly typical of internet fora and the like. Still the phenomenon is, in our view, remarkable not only for the history of the Persian language, but potentially also for grammaticalisation studies.⁵⁴

‘Le renouvellement incessant du vocabulaire tient à l’essence de la langue, à sa fonction de moyen de communication entre les membres de la société, ce qui la relie directement à toute activité de l’homme dans toutes les sphères de son travail’ Telegdi (1951: 143).

ACKNOWLEDGEMENTS

We would like to thank Pegah Ebrahimi, Maryam Mokhtari Afjedi and Hamed Saneey for their invaluable assistance in the development of the questionnaires as well as in obtaining data. Our thanks also go to all the speakers who replied to the questionnaires or replied to various enquiries, and to the participants of the conference ICIL 9 (August 2021) for comments and questions. We are very grateful to Thomas Jügel for commenting on a previous version as well as to two anonymous reviewers for their constructive feedback. Murad Suleymanov kindly suggested colloquial English translations for the examples. We are also grateful to Nicholas Sims-Williams for advice and assistance. [p. 47]

⁵² Another parallelism is the semantic bleaching characteristic for grammaticalisation processes, which can in fact be also demonstrated for Persian light verbs (Korn 2013).

⁵³ See also Haspelmath & Michaelis (2017), who mention ‘strong general tendencies toward analyticization in language change’ (p. 9) and note that ‘a diachronic process of ‘analyticizing’ or ‘refunctionalizing’ is widespread and is involved in a substantial number of salient grammatical innovations’ (p. 4).

⁵⁴ Needless to say, Persian is not the only language showing phenomena disagreeing with the cycle. Romanian employs analytic constructions in the verbal domain for expressing various TAM categories (like other Romance languages such as French), but, as pointed out by Dragomirescu et al. (2022: 243), ‘Old Romanian is characterized by a larger number of analytic verb forms than modern Romanian’. However, while the loss of analytic patterns (with certain synthetic ones being preserved) may be unexpected, the phenomenon is different from the development (or revival) of new synthetic structure found in Persian.

ABBREVIATIONS⁵⁵

1, 2, 3: first, second, third person	PL: plural
COP: copula	PN: name
DEM: demonstrative pronoun	PRF: perfect (participle)
EZ: ezafe	PROG: progressive
IDV: individuation marker	PRS: present (stem)
IMP: imperative mood	PST: past (stem)
IMPF: imperfect	PTCP: participle
INF: infinitive	REF: referential
IPFV: imperfective	SBJV: subjunctive mood
itr.: intransitive	SG: singular
NEG: negation	SUB: subordinator
PC: pronominal clitic (enclitic pronoun)	tr.: transitive

Correspondence

Dorian Pastor

École Normale Supérieure / Université PSL

*Email: dorian.pastor@ens.psl.eu

REFERENCES

- BAAYEN, R. HARALD, 2008. *Analyzing Linguistic Data. A Practical Introduction to Statistics*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511801686>
- BATENI, MOHAMMAD REZĀ, 1990. 'Fārsi zabān-i 'aqim? [Persian, a sterile language?]', *Adineh* 33. 66–71.
- 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*. New York: Cambridge University Press. 48–78.
<https://doi.org/10.1017/CBO9780511712234.004>
- CIANCAGLINI, CLAUDIA, 2011. 'The formation of the periphrastic verbs in Persian and neighbouring languages', in Mauro Maggi & Paola Orsatti (eds.), *The Persian Language in History*. Wiesbaden: Reichert. 3–21.
- DEHKODĀ, 'ALI AKBAR, 1399/2020. *Loġat-nāme-ye Dehxodā (nosxe-ye dijital), bar asās-e nosxe-ye fiziki-ye 15 jeldi, entešār-e sāl-e 1377. Mo'assese-ye loġat-nāme-ye Dehxodā va markaz-e beyn ol-melali-e āmuzeš-e zabān-e Fārsi, dānešgāh-e Tehrān*. [Dehkhoda Dictionary (Digital version), based on the physical 15-volume version, published in 1998. Dehkhoda Dictionary Institute and International Center for Persian Language Teaching, University of Tehran].
<http://dehkhoda.ut.ac.ir>
- DIXON, ROBERT, 1997. *The Rise and Fall of Languages*. Cambridge/New York: Cambridge University Press. <https://doi.org/10.1017/CBO9780511612060>
- DRAGOMIRESCU, ADINA, ALEXANDRU NICOLAE & RODICA ZAFIU, 2022. 'The loss of analyticity in the history of Romanian verbal morphology', in Adam Ledgeway, John Charles Smith and Nigel Vincent (eds.), *Periphrasis and Inflexion in Diachrony. A View from Romance*. Oxford: Oxford University Press. 241–271. <https://doi.org/10.1093/oso/9780198870807.003.0010>
- GHOMESHI, JILA, 1997. 'Topics in Persian VPs', *Lingua* 102(2). 133–167.
[https://doi.org/10.1016/S0024-3841\(97\)00005-3](https://doi.org/10.1016/S0024-3841(97)00005-3) [p. 48]
- HASPELMATH, MARTIN & SUSANNE MICHAELIS, 2017. 'Analytic and synthetic: Typological change in varieties of European languages', in Isabelle Buchstaller & Beat Siebenhaar (eds.), *Language Variation - European Perspectives VI. Selected papers from the Eighth International Conference on Language Variation in Europe (ICLaVE 8), Leipzig, May 2015*. Amsterdam: Benjamins. 3–22. <https://doi.org/10.1075/silv.19.01has>

⁵⁵ For the sources of the examples, see note 11.

- HAYYIM, SULEYMAN, 1933-1935. *New Persian-English Dictionary, Complete and Modern, Designed to Give the English Meanings of Over 50,000 Words, Terms, Idioms, and Proverbs in the Persian language, As well as the Transliteration of the Words in English Characters. Together with a Sufficient Treatment of All the Grammatical Features of the Persian Language*. Tehran: Librairie-imprimerie Bérroukhim. <https://dsal.uchicago.edu/dictionaries/hayyim/>
- KARIMI, SIMIN & RYAN WALTER SMITH, 2020. ‘Another look at Persian *rā*’, in Richard K. Larson, Sedigheh Moradi & Vida Samiian (eds.), *Advances in Iranian Linguistics*. Amsterdam / Philadelphia: Benjamins. 155–172. <https://doi.org/10.1075/cilt.351.09kar>
- KORN, AGNES, 2009. ‘Lengthening of *i* and *u* in Persian’, in Werner Sundermann, Almut Hintze & François de Blois (eds.), *Exegisti Monumenta. Festschrift in Honour of Nicholas Sims-Williams*. Wiesbaden: Harrassowitz. 197–213.
- KORN, AGNES, 2013. ‘Looking for the middle way: Voice and transitivity in complex predicates in Iranian’, *Lingua* 135. 33–55. <https://doi.org/10.1016/j.lingua.2013.07.015>
- LABOV, WILLIAM, 1990. ‘The intersection of sex and social class in the course of linguistic change’, *Language Variation and Change* 2(2). 205–254.
- LAZARD, GILBERT, 2006. *Grammaire du persan contemporain, Avec la collaboration de Yann Richard, Rokhsareh Hechmati et Pollet Samvelian, nouvelle édition*. Tehran: Institut Français de Recherche en Iran.
- MIESTAMO, MATTI, 2005. *Standard Negation: The Negation of Declarative Verbal Main Clauses in a Typological Perspective*. Berlin/Boston: De Gruyter. <https://doi.org/10.1515/9783110197631>
- MIESTAMO, MATTI & JOHAN VAN DER AUWERA, 2011. ‘Negation and perfective vs. imperfective aspect’, in Jesse Mortelmans, Tanja Mortelmans & Walter De Mulder (eds.), *From Now to Eternity. Cahiers Chronos 22*. Amsterdam: Rodopi. 65–84. https://doi.org/10.1163/9789042032682_005
- MIR-DJALALI, ELAHE, 1974. *Recherche sur la structure transformationnelle du verbe persan : le verbe simple*. Paris: University of Paris VII, dissertation.
- NÆSS, ÅSHILD, 2007. *Prototypical Transitivity*, Amsterdam/ Philadelphia: Benjamins. <https://doi.org/10.1075/tsl.72>
- STEINGASS, FRANCIS J., 1892. *A Comprehensive Persian-English Dictionary, Including the Arabic Words and Phrases to be Met With in Persian Literature*. Routledge & K. Paul, London. <https://dsal.uchicago.edu/dictionaries/steingass/>
- TELEGDI, ZSIGMOND, 1951. ‘Nature et fonction des périphrases verbales dites “verbes composés” en persan’, *Acta Orientalia* 2(3). 315–338. <https://www.jstor.org/stable/23656236>
- TRUDGILL, PETER, 1972. ‘Sex, covert prestige and linguistic change in the urban British English of Norwich’, *Language in Society* 1(2). 179–195. <https://doi.org/10.1017/S0047404500000488>
- VAFAEIAN, GHAZALEH, 2018. *Progressives in Use and Contact: A Descriptive, Areal and Typological Study With Special Focus on Selected Iranian Languages*, Stockholm: Stockholm University. <http://su.diva-portal.org/smash/record.jsf?pid=diva2%3A1200741&dswid=675>
- VAN GELDEREN, ELLY, 2011. *The Linguistic Cycle – Language Change and the Language Faculty - Oxford Scholarship*, New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199756056.001.0001>
- WATKINS, CALVERT, 1962. *Indo-European Origins of the Celtic Verb. I. The Sigmatic Aorist*. Dublin: Dublin Institute for Advanced Studies.
- WINDFUHR, GERNOT & JOHN R. PERRY. 2009. ‘Persian and Tajik’, in Gernot Windfuhr (ed.), *The Iranian Languages*. London/New York: Routledge. 416–544.